Ultrasound imaging using coded signals

Misaridis, Athanasios

Publication date:
2001

Document Version
Publisher's PDF, also known as Version of record

Thanassis Misaridis

Ultrasound Imaging Using Coded Signals

Center for Fast Ultrasound Imaging

Technical University of Denmark

AUGUST 2001

Ultrasound Imaging Using Coded Signals

Thanassis Misaridis

Center for Fast Ultrasound Imaging
Technical University of Denmark

August 2001

SUBMITTED IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY
AT
THE TECHNICAL UNIVERSITY OF DENMARK
AUGUST 2001

_____
Signature of Author

To my parents . . .
who—no matter what—
have been the global constant
of my life

Contents

Preface

The Ph.D. project entitled ”Ultrasound imaging using coded signals” has been part of the research activities of the Center For Fast Ultrasound Imaging (CFU), directed by my advisor Prof. Jørgen Arendt Jensen during the years 1998-2001. The Center is funded by the Danish Research Council, the Danish ultrasound manufacturer B-K Medical and Herlev University Hospital in Denmark. I would like to thank all contributors for their financial support.

During the three years of the Ph.D. project, the following papers related to coded excitation have been published:

- T. X. Misaridis and K. Gammelmark and C. H. Jørgensen and N. Lindberg and A. H. Thomsen and M. H. Pedersen and J. A. Jensen. Potential of coded excitation in medical ultrasound imaging. *Ultrasonics*, 38:183–189, 2000.
- T. X. Misaridis and J. A. Jensen. An effective coded excitation scheme based on a predistorted FM signal and an optimized digital filter. In *Proc. IEEE Ultrason. Symp.*, volume 2, pages 1589–1593, 1999.
- T. X. Misaridis, M. H. Pedersen, and J. A. Jensen. Clinical use and evaluation of coded excitation in B-mode images. In *Proc. IEEE Ultrason. Symp.*, volume 2, pages 1689–1693, 2000.
- T. Misaridis and J. A. Jensen. Space-time encoding for high frame rate ultrasound imaging. To be published in *Ultrasonics*, 2002.

The impetus for this research project has been the increasing interest over the last decade in the medical ultrasound community in the utilization of more sophisticated excitation signals than the single-carrier short pulses currently used in ultrasound scanners. The potential advantages of such ”coded” signals are two: a) an increase in penetration depth and/or an increase in signal-to-noise ratio (SNR), and b) an increase in frame rate. Both signal-to-noise ratio and frame rate are very valuable resources in medical ultrasound imaging. Higher SNR will allow imaging of structures that are located deep inside the human body. Higher SNR can also allow migration to higher frequencies, which in turn will result in images with better resolution. High frame rates will make

real-time three-dimensional ultrasound imaging possible and will allow imaging of fast moving objects such as the heart.

Coded signals have been used successfully in other engineering disciplines such as radars and mobile communication systems. It is therefore natural for one to ask for the reasons why coded excitation has not been explored and used in medical ultrasound imaging as much as in the other areas. The answer to this question (apart from the required complexity in electronics and implementation issues) is that ultrasound imaging with codes is a far more challenging and difficult task. In radar systems, the problem is the detection of isolated targets. In imaging, the problem is mapping of distributed scatterers, where no decision-making is possible. The high requirements in the displayed dynamic range of the ultrasound images is translated to increased requirements for the correlation properties of the coded signals. The problem is further complicated by the frequency-dependent attenuation in the tissues and by the presence of speckle. In communication systems, codes are used as modulated carriers of binary data and separation of users is based on threshold detectors. For fast ultrasound imaging, any cross-talk between simultaneously transmitted coded beams will appear as ghost echoes in the image. Apart from having a more difficult task to accomplish, the ultrasound engineer has to work with far more limited system bandwidth and code length. Unfortunately, the performance of coded excitation is based exactly on these two parameters.

The aim of this dissertation is to investigate systematically the applicability of modulated signals in medical ultrasound imaging and to suggest appropriate methods for coded imaging. This book is an attempt to provide to the ultrasound community with an overview of the problems, possibilites and expected benefits from application of modulated signals in ultrasound imaging. The author hopes that the principles and ideas presented and discussed here will inspire others in designing coded imaging systems in the future with improved performance.

Abstract

Modulated (or coded) excitation signals can potentially improve the quality and increase the frame rate in medical ultrasound scanners. The aim of this dissertation is to investigate systematically the applicability of modulated signals in medical ultrasound imaging and to suggest appropriate methods for coded imaging, with the goal of making better anatomic and flow images and three-dimensional images. On the first stage, it investigates techniques for doing high-resolution coded imaging with improved signal-to-noise ratio compared to conventional imaging. Subsequently it investigates how coded excitation can be used for increasing the frame rate. The work includes both simulated results using Field II, and experimental results based on measurements on phantoms as well as clinical images.

Initially a mathematical foundation of signal modulation is given. Pulse compression based on matched filtering is discussed. Correlation and compression properties of coded signals are shown to depend on a single parameter of the coded signals: the time-bandwidth product. It is shown that, due to attenuation in the tissues, the matched flter output is related to the ambiguity function of the excitation signal. Although a gain in signal-to-noise ratio of about 20 dB is theoretically possible for the time-bandwidth product available in ultrasound, it is shown that the effects of transducer weighting and tissue attenuation reduce the maximum gain at 10 dB for robust compression with low sidelobes.

Frequency modulation and phase modulation are considered separately and their resolution, side-lobes, expected signal-to-noise gain and performance in tissue imaging are discussed in detail. A method to achieve low compression sidelobes by reducing the ripples of the amplitude spectrum of the FM signals is described.

Application of coded excitation in array imaging is evaluated through simulations in Field II. The low degree of the orthogonality among coded signals for ultrasound systems is first discussed, and the effect of mismatched filtering in the cross-correlation properties of the signals is evaluated. In linear array imaging it is found that the frame rate can be doubled without any degradation in image quality, by using two coded sequences that have a cross-correlation of at least 11 dB. Other coding schemes that can increase the frame rate by nearly 5 times with a small compromise in resolution are discussed. Coded synthetic transmit aperture imaging with only 4 emissions is shown to yield the same signal-to-noise ratio as with conventional phased-array imaging which

uses 51 emissions. Further frequency-division coding can make it possible to obtain images with acceptable resolution with only two emissions. Finally, a novel coding technique which uses pulse train excitation is presented.

Acknowledgements

I would like to thank:

- My advisor **Prof. Jørgen Arendt Jensen** for being the inspired scientist he is, and for simply being the ideal person to work for. I would like to thank him for his support, for making work a pleasure by infusing to me some of his professionalism, enthusiasm and visions, for letting me work my own schedule, yet always there to answer my questions, and for the innumerous things I have learnt from him, everything from Linux, signal processing and ultrasound imaging to organization skills and scientific ethics.
- **Dr. Peter Munk** for following the research progress closely throughout the project, for teaching me a great deal about ultrasound imaging, for making invaluable comments and giving ideas and research directions, for helping me with hardware issues (and often doing my work...), for always being supportive, helpful and discrete, and for a hundred more reasons.
- Ph.D. student **Borislav Tomov** for his great help on hardware issues, and his extensive tests and scripts he provided me with for the experimental system. I would especially like to thank him for always being there and helping me throughout the experimental system setup.
- Ph.D. student **Svetoslav Nikolov** for writing his beam formation Matlab toolbox that has decreased the simulation time significantly, for writing the software for the experimental system, for various other Matlab scripts he has provided me with, for useful advice on Linux issues, for useful discussions and exchange of ideas on beamforming and imaging in general, and lastly for making a great office neighbor.
- System administrator **Henrik Laursen** for his great support on Linux and network issues and his kindness to help me setting up my laptop.
- Technician **Finn Pedersen** for building the interface box to the scanner.
- Students **Kim Gammelmark, Christian H. Jørgensen, Niklas Lindberg** and **Anders H. Thomsen** for their valuable work on the experimental setup and acquiring several clinical images.

- M.D. **Morten Pedersen** for some clinical scans.
- **Ellen Nagato Watanabe** for her patience and support as my girlfriend during these hard-working years.
- And last but not least, graphic designer **Maria Candia** for designing the hard cover of this dissertation, for a wonderful painting and for her great support during the last months.

Nomenclature

Symbols

$a(t)$	Amplitude modulation function
B	Signal bandwidth
B_r	Relative bandwidth for Gaussian pulse
β	rms signal bandwidth
γ	Mismatch factor for FM signals
$\gamma(t)$	Complex matched filter output
$\gamma_w(t)$	Complex mismatched filter output
c	Speed of sound
D	Time-bandwidth product ($D = TB$)
δ	rms signal duration
E	Signal energy
f_i	Instantaneous frequency
f_0	Transducer center frequency
f_d	Frequency downshift of the received signal
f_m	Frequency downshift parameter of the matched filter
f_{prf}	Pulse repetition frequency
f_s	Sampling frequency
$g(t)$	Real matched filter output
$H(f)$	Real matched filter transfer function
$\mathrm{H}(f)$	Complex matched filter transfer function
$\mathrm{H}_w(f)$	Complex mismatched filter transfer function
$h(t)$	Real matched filter impulse response
$\eta(t)$	Complex matched filter impulse response
I_m	Maximum intensity
I_{sppa}	Spatial peak, pulse average intensity
I_{spta}	Spatial peak, temporal average intensity
λ	Wavelength
$M(f)$	Fourier transform of $\mu(t)$
μ	FM slope (FM sweep rate)

$\mu(t)$	Complex envelope of $\psi(t)$
N	Number of pulses in a pulse train
$N(f)$	Noise power density
$N_c(f)$	Speckle power density spectrum
$r(t)$	Received signal
$R_{ss}(\tau)$	Auto-correlation function of $s(t)$
$R_{\psi\psi}(\tau)$	Auto-correlation function of $\psi(t)$
SNR	Signal-to-noise ratio
$S(f)$	Fourier transform of $s(t)$
$s(t)$	Real modulated signal
σ	Standard deviation
σ^2	Variance
t	Time
T	Signal duration
τ	Lag in correlation function
$\tau_g(f)$	Group delay function
$\phi(t)$	Phase modulation function
$\Psi(f)$	Fourier transform of $\psi(t)$
$\chi(\tau, f_d)$	Ambiguity function
$\chi_{nm}(\tau, f_d)$	Cross-ambiguity function
$\psi(t)$	Complex modulated signal
z	Depth (axial distance from transducer surface)

Abbreviations

AC	Auto-correlation
ACF	Auto-correlation function
AF	Ambiguity function
AFG	Arbitrary function generator
BP	Band-pass
B-mode	Brightness mode
CAF	Cross-ambiguity function
CC	Cross-correlation
CCF	Cross-correlation function
CFM	Color flow mapping
CW	Continuous wave
FFT	Fast Fourier transform
FM	Frequency modulation
FIR	Finite impulse response
FSK	Frequency shift keying
GSNR	Gain in signal to noise ratio
ISL	Integrated sidelobe level
LP	Low-pass
MSR	Mainlobe-to-sidelobe ratio

NLFM	Non-linear frequency modulation
PM	Phase modulation
PN	Pseudo-noise (sequences)
PSF	Point spread function
PSK	Phase shift keying
PSL	Peak sidelobe level
PW	Pulsed wave
QLFM	Quantized linear frequency modulation
RF	Radio frequency
RMS	Root mean square
ROI	Region of interest
SNR	Signal to noise ratio
STA	Synthetic transmit aperture
TB	Time-bandwidth product
TGC	Time gain compensation

List of Figures

List of Tables

CHAPTER

ZERO

Introduction

The use of ultrasound in medical diagnostic applications has a history of half a century. There are several books published on ultrasound physics, the principles of image formation and beamforming in ultrasound imaging. For more information, the reader is referred to the relevant literature [2, 4, 5, 6, 7, 8, 9].

0.1 Potential advantages of coded excitation

In ultrasound imaging, signal-to-noise ratio (SNR) is a crucial factor for image quality. The severe attenuation of the ultrasonic signals in the tissue results in echoes from large depths literally buried in noise. Flow estimation or synthetic aperture techniques are two of the fields in ultrasound imaging that suffer the most from the low SNR. On the other hand, resolution requirements favor transmission of short pulses, and thereby low signal energy. The transmitted power should then be raised proportionally to the shortening of the pulse. Unfortunately, the peak intensity levels that are permitted by the FDA (Food and Drug Administration) to be sent into the human body set a lower limit in pulse duration. Transmission of modulated signals can improve the SNR a great deal without degrading imaging resolution. This is achieved by retaining the system bandwidth without reducing the pulse width.

Current ultrasound scanners form the ultrasonic image by emitting a pulsed field in one direction. The scattered field is then received and focused in the same direction. This is repeated for a number of lines, in order to assemble an image. The frame rate is, thus, limited by the speed of sound in tissue and the number of directions for an image. In blood flow imaging, a number of pulses must be emitted in one direction in order to estimate the velocity, and this can -in some investigations over large regions of interest- lower the frame rate to an unacceptable rate. The number of lines in one direction for flow images also determines the accuracy and bias of the velocity estimates, since the removal of stationary echoes and the subsequent velocity estimate are improved proportionaly to the number of lines. The sequential acquisition limits the number of lines per second, and makes

it impossible to obtain real-time 3-D ultrasound images, as lines here have to be acquired for a full volume instead of a cross-sectional image. Many problems can be solved and advantages gained by increasing the possible acquisition rate. Modulated signals provide a large waveform diversity, with the potential of increasing the frame rate in ultrasound imaging. How this can be done is the topic of the second part of the dissertation.

0.2 Literature Review

The first investigator that considered the application of coded excitation in medical ultrasound systems was Takeuchi [10] in a paper dating back in 1979. Takeuchi pointed out the time-bandwidth limitations in the application of coded signals in ultrasound imaging. Possibly due to this limitation, as well as the anticipated limitation imposed by the frequency-dependent attenuation in tissues, there is no much contribution during the following years in the literature on this topic. It is only in the last decade, that there is a renewed interest in coded excitation within the medical ultrasound community, resulting in a rather vast amount of published papers.

O' Donnell [11] discussed the expected improvement in signal-to-noise ratio, concluding that coded excitation can potentially yield an improvement of 15 to 20 dB. His system was using a single correlator on the output of a digital beamformer, i.e. beamforming was done prior to compression. This approach, although advantageous in terms of implementation, poses requirements on the code length and arises issues about the effect of time-gain compensation (TGC) and dynamic focusing on pulse compression.

Subsequently, several contributions have been made, primarily on pulse compression mechanism and sidelobe reduction problems. Considerably less authors [12, 13, 14] have considered fast ultrasound imaging using coded signals with low cross-correlation properties. Some authors [15, 16, 11] have considered the application of inverse filtering instead of matched filtering for more efficient sidelobe reduction. Most of the authors have used chirp (linear FM) or pseudo-chirp excitation [11, 17, 18, 19, 20], others have considered binary codes, such as m-sequences [16, 14] and orthogonal Golay sequences [12], and others have considered both [15]. Rao [19] pointed out that ultrasonic attenuation will result in SNR degradation. Pollakowski and Ermert [18] discussed the design of non-linear FM signals. The same group [21] has also considered frequency-dependent filtering in order to compensate for the attenuation.

0.3 Thesis structure

The dissertation is organized as following:

First, the mathematical properties of coded signals and pulse compression are discussed, with the ultrasound-specific requirements in mind. Subsequently FM- and phase-coded signals are discussed and appropriate coded signals, weighting and mismatched filtering are designed. Imaging with single coded excitation is simulated and tested in clinical images, and the findings in terms of SNR improvement, robustness and resolution are discussed. This concludes the first part of the disseration (chapters 1 to 7).

The second part (chapters 8 to 10) is devoted to the usefulness of coded signals as a means of increasing the frame rate in ultrasound imaging. It discusses cross-correlation and orthogonality properties among sets of coded signals and fast coded imaging techniques in linear-array imaging, synthetic aperture imaging and pulse train imaging.

- **Chapter 1** gives a brief introduction on modulated signal representations and basic signal properties concepts.
- **Chapter 2** describes the matched filter, the center element of pulse compression and the ambiguity function associated with coded waveforms. Emphasis is given on the application in ultrasound imaging.
- **Chapter 3** presents the frequency-modulated (FM) signals.
- **Chapter 4** describes filtering techniques for sidelobe reduction using FM signals.
- **Chapter 5** introduces phase-modulated signals. Compression properties of binary phase as well as polyphase coded signals are presented. Complementary codes and orthogonal Hadamard matrices are also discussed.
- **Chapter 6** discusses the application of coded excitation in ultrasound imaging, as well as measures of resolution and SNR improvement.
- **Chapter 7** presents experimental results of coded excitation in ultrasound imaging. Phantom and clinical images are shown.
- **Chapter 8** examines the cross-correlation properties of FM signals and phase-coded signals.
- **Chapter 9** presents simulated results in coded ultrasound imaging with high frame rate. Linear array imaging and synthetic transmit aperture imaging are discussed.
- **Chapter 10** discusses the potential of a novel coding technique using acoustically generated pulse trains.
- **Chapter 11** summarizes briefly the findings of the dissertation.

CHAPTER

ONE

Modulated signals

1.1 Introduction

This chapter gives a rather quick overview of basic concepts in signal analysis. A single measure of a signal - the time-bandwidth (TB) product - will be used in order to show the equivalence between the terms modulated signal, high time-bandwidth product, and pulse compression.

1.2 Signal basics

Let a real modulated signal $s(t)$ be expressed as:

$$s(t) = a(t) \cdot \cos\left[2\pi f_0 t + \varphi(t)\right], \tag{1.1}$$

where

$a(t)$ is the amplitude modulation function and

$\varphi(t)$ is the phase modulation function.

The argument of the cosine in (1.1) is the phase function $\Phi(t)$ of the signal :

$$\Phi(t) = 2\pi f_0 t + \varphi(t) \tag{1.2}$$

If $\varphi(t)$ is a continuous time function, the time derivative of the phase is defined as the *instantaneous frequency* f_i:

$$f_i = \frac{1}{2\pi}\frac{d\Phi(t)}{dt} = f_0 + \frac{1}{2\pi}\frac{d\varphi(t)}{dt} \tag{1.3}$$

From (1.2) it can be seen that the phase modulation function has to be a non-linear function of time, since any linear term can be combined with the carrier frequency. If the amplitude $a(t)$ varies

slowly compared to the instantaneous frequency f_i, $|a(t)|$ represents essentially the envelope of the signal.

The Fourier transform of the signal $s(t)$ is denoted as $S(f)$. $s(t)$ and $S(f)$ are related through the Fourier integrals:

$$\begin{aligned} S(f) &= \int_{-\infty}^{\infty} s(t) \cdot e^{-j2\pi ft} dt \\ s(t) &= \int_{-\infty}^{\infty} S(f) \cdot e^{j2\pi ft} df \end{aligned} \tag{1.4}$$

The energy of the signal E is given by:

$$E = \int_{-\infty}^{\infty} [s(t)]^2 dt = \int_{-\infty}^{\infty} |S(f)|^2 df \tag{1.5}$$

where the second part of the equation is obtained by Parseval's theorem.

Substituting (1.1) in (1.5) we obtain

$$E = \frac{1}{2} \int_{-\infty}^{\infty} [a(t)]^2 dt + \frac{1}{2} \int_{-\infty}^{\infty} [a(t)]^2 \cos\{2[2\pi f_0 t + \varphi(t)]\} dt \tag{1.6}$$

For narrowband signals, the frequencies contained in the functions $a(t)$ and $\varphi(t)$ are small compared to the carrier frequency f_0. In this case, the second integral represents the oscillations of a sine under a slowly varying envelope and is essentially zero. Then, the energy can be approximated by:

$$E \approx \frac{1}{2} \int_{-\infty}^{\infty} [a(t)]^2 dt \tag{1.7}$$

This result shows that as long as the phase modulation does not distort the signal envelope, the signal energy is not altered.

The auto-correlation function is defined by the integral:

$$R_{ss}(\tau) = \int_{-\infty}^{\infty} s(t)s(t-\tau) dt = \int_{-\infty}^{\infty} |S(f)|^2 e^{j2\pi f\tau} df \tag{1.8}$$

The auto-correlation shows how different a signal is compared to its shifted versions as a function of the time shift τ. The maximum occurs when $\tau = 0$, and is equal to the signal energy:

$$R_{ss,\max}|_{\tau=0} = \int_{-\infty}^{\infty} [s(t)]^2 dt = E \tag{1.9}$$

1.3 Complex notation of narrowband signals

Signals used in practice are real, however the complex notation offers many advantages particularly in expressing correlation integrals. The matched filter response -the core of the receive processing in coded excitation systems - is a correlation integral and therefore the complex notation is very convenient.

Since the spectrum of a real signal is symmetric around the zero frequency, an equivalent but simplified notation is a complex signal that has no negative frequencies and double the amplitude of the positive frequencies. A complex signal is called *analytic* if its spectrum consist of only positive frequencies. This is possible when the real and imaginary parts of the signal form a Hilbert pair [22]. Let

$$\psi(t) = \mu(t) \cdot e^{j2\pi f_0 t}. \tag{1.10}$$

be an analytical signal, whose real part is equal to the modulated signal given in (1.1). $\mu(t)$ is a complex function, usually referred to as the *complex envelope* [1], and combines amplitude and phase modulation:

$$\mu(t) = |\mu(t)| \cdot e^{j\phi(t)}. \tag{1.11}$$

The real waveform is derived as the real part of the complex signal:

$$s(t) = \mathrm{Re}\{\psi(t)\} = |\mu(t)| \cdot \cos[2\pi f_0 t + \phi(t)] \tag{1.12}$$

If $\Psi(f)$ and $M(f)$ are the Fourier transforms of the analytic signal $\psi(t)$ and the complex envelope $\mu(t)$ respectively, the Fourier transform of (1.10) yields

$$M(f) = \Psi(f + f_0). \tag{1.13}$$

Thus, the frequency spectrum of the complex envelope is the shifted spectrum of the signal with the carrier frequency removed. When the real signal is narrowband the conditions

$$a(t) \approx |\mu(t)| \quad , \quad \varphi(t) \approx \phi(t) \tag{1.14}$$

are satisfied and the analytic complex signal is derived from the real signal simply by substituting the cosine with an exponent. Note that the resulting signal (sometimes referred to as the *exponential* signal) will not be strictly analytic, if the fractional bandwidth of the real signal is so high, that the spectrum of the exponential signal folds over the negative frequencies.

Using the second part of (1.5) and the fact that $\Psi(f) = 2S(f)$ for positive frequencies, the energy can now be written as [1]:

$$\begin{aligned} E &= \int_{-\infty}^{\infty} |S(f)|^2 df = \int_{0}^{\infty} \left|\tfrac{1}{2}\Psi(f)\right|^2 df = \tfrac{1}{2}\int_{-\infty}^{\infty} |\Psi(f)|^2 df \\ &= \tfrac{1}{2}\int_{-\infty}^{\infty} [\psi(t)]^2 dt = \tfrac{1}{2}\int_{-\infty}^{\infty} |\mu(t)|^2 dt = \tfrac{1}{2}\int_{-\infty}^{\infty} |M(f)|^2 df \end{aligned} \tag{1.15}$$

The equality sign in the latter part of (1.15) is exact, as opposed to the approximation in (1.7). This is another indication that going from the real to the complex notation is only an approximation [1]. In the rest of the analysis, it is assumed that the exponential signal is a good approximation of the analytic signal, an assumption that is reasonable for the relatively narrowband signals that can pass from an ultrasound transducer.

1.4 Correlation integrals

In a similar manner to the definition given in (1.8), the complex auto-correlation of $\psi(t)$ is given by:

$$R_{\psi\psi}(\tau) = \int_{-\infty}^{\infty} \psi(t)\psi^*(t-\tau)dt = \int_{-\infty}^{\infty} |\Psi(f)|^2 e^{j2\pi f\tau} df. \tag{1.16}$$

Using (1.10) and (1.13), the auto-correlation can be expressed as a function of the modulation function:

$$R_{\psi\psi}(\tau) = e^{j2\pi f_0\tau} \int_{-\infty}^{\infty} \mu(t)\mu^*(t-\tau)dt = e^{j2\pi f_0\tau} \int_{-\infty}^{\infty} |M(f)|^2 e^{j2\pi f\tau} df \tag{1.17}$$

The analytic signal of $s(t)$ is

$$\psi(t) = s(t) + jH\{s(t)\}, \tag{1.18}$$

where H denotes the Hilbert transform. Using (1.10) and the symmetry properties of the Hilbert transform, the auto-correlation function of $\psi(t)$ can be found [23]:

$$R_{\psi\psi}(\tau) = 2\left[R_{ss}(\tau) + j \cdot H\{R_{ss}(\tau)\}\right] \tag{1.19}$$

The above equation can also be derived by taking the real part of (1.16) and use the one-sided spectrum property of $\psi(t)$.

In an imaging system, the displayed quantity is the envelope of the signal. As it will be shown in the next chapter, the envelope of the real auto-correlation function is the matched-filter response, and is in fact the inverse Fourier transform of the modulation's energy density spectrum $|M(f)|^2$:

$$Env\{R_{ss}(\tau)\} = \frac{1}{2}\left|R_{\psi\psi}(\tau)\right| = \frac{1}{2}\left|\int_{-\infty}^{\infty} \mu(t)\mu^*(t-\tau)dt\right| = \frac{1}{2}\left|\int_{-\infty}^{\infty} |M(f)|^2 e^{j2\pi f\tau} df\right| \tag{1.20}$$

Fig. 1.1 sketches the application of (1.20) in the estimate of the auto-correlation envelope of a single-carrier pulse of length T. In the absence of modulation, $\mu(t)$ is a real-valued rectangular

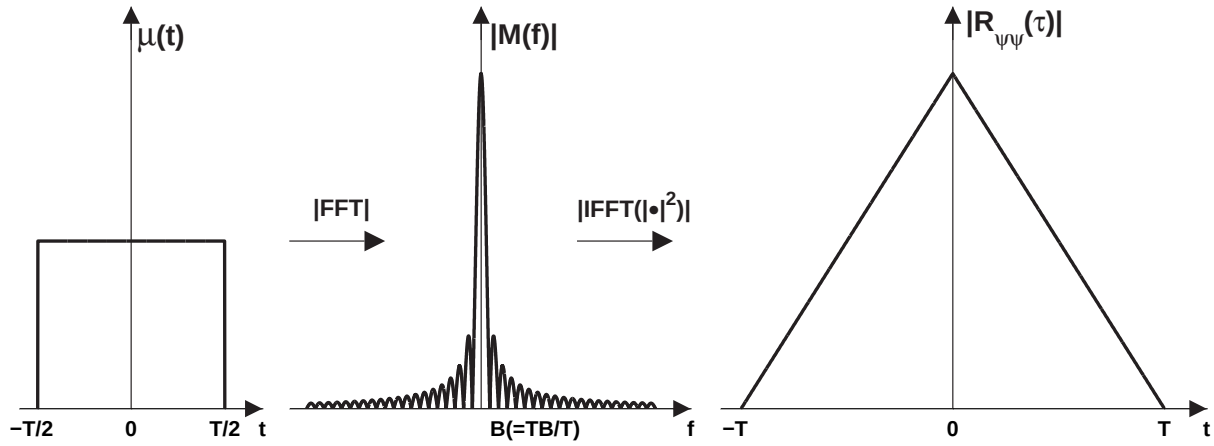

Figure 1.1: Triangular auto-correlation envelope of a constant-carrier pulse using (1.20).

window. The modulus of its Fourier transform $M(f)$ is a *sinc* function, and the inverse Fourier transform of a $sinc^2$ function is the triangle function.

Fig. 1.2 illustrates the envelope of the auto-correlation function for a linear FM signal with a time-bandwidth product of 20.

1.5 Waveform parameters and the uncertainty principle

Rihaczek [1] has defined waveform parameters for the effective signal duration and bandwidth as the second moments of the energy signals as following:

The *rms signal duration* δ:

$$\delta^2 = (2\pi)^2 \frac{\int_{-\infty}^{\infty} t^2 |\psi(t)|^2 dt}{\int_{-\infty}^{\infty} |\psi(t)|^2 dt} = \left(\frac{2\pi}{\sqrt{2E}}\right)^2 \int_{-\infty}^{\infty} t^2 |\psi(t)|^2 dt = \left(\frac{2\pi}{\sqrt{2E}}\right)^2 \int_{-\infty}^{\infty} t^2 |\mu(t)|^2 dt, \qquad (1.21)$$

and the *rms signal bandwidth* β:

$$\beta^2 = (2\pi)^2 \frac{\int_{-\infty}^{\infty} (f - f_0)^2 |\Psi(f)|^2 df}{\int_{-\infty}^{\infty} |\Psi(f)|^2 df} = \left(\frac{2\pi}{\sqrt{2E}}\right)^2 \int_{-\infty}^{\infty} f^2 |M(f)|^2 df. \qquad (1.22)$$

Using the Fourier relationship:

$$\mu'(t) \overset{F}{\longleftrightarrow} j2\pi f \cdot M(f), \qquad (1.23)$$

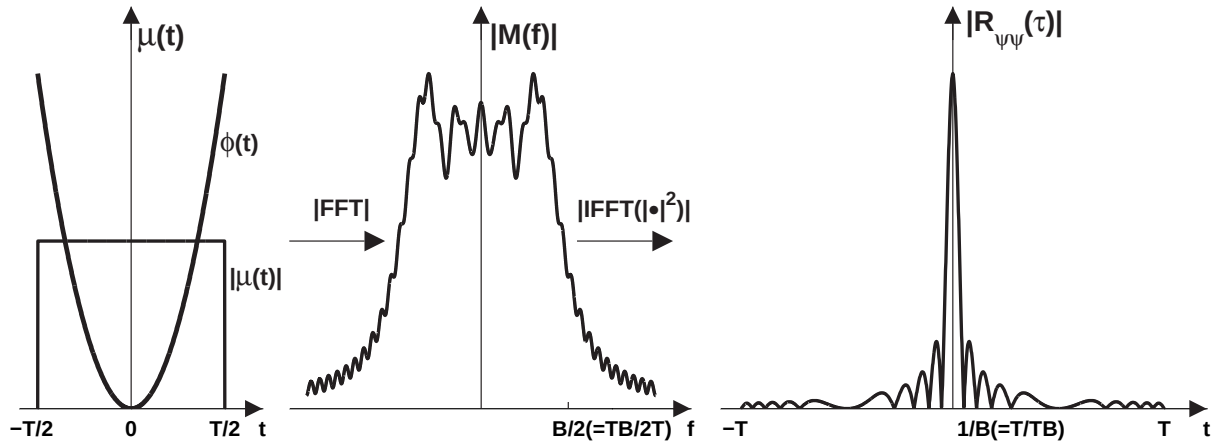

Figure 1.2: Application of (1.20) in the estimation of the auto-correlation envelope for a linear FM signal with a time-bandwidth product of 20. The modulation function $\mu(t)$ has a rectangular envelope and a quadratic phase. As it will be derived in Chapter 4, this signal has an approximate rectangular amplitude spectrum. The auto-correlation envelope is the inverse Fourier transform of a rectangular, i.e. approximately a *sinc* function.

the rms bandwidth can be expressed alternatively as a function of the derivative of the modulation function:

$$\beta^2 = \frac{1}{2E}\int_{-\infty}^{\infty} \left|\mu'(t)\right|^2 dt \tag{1.24}$$

The advantage of these generic definitions over the conventional definitions of the signal duration T and bandwidth B is that they do not depend on the time envelope and the shape of the spectrum respectively. However, they offer a sufficient description of the waveform properties in single parameters.

Multiplying (1.21) with (1.24) the time-bandwidth product of any waveform is:

$$\delta^2\beta^2 = \left(\frac{2\pi}{2E}\right)^2 \int_{-\infty}^{\infty} t^2 \left|\mu(t)\right|^2 dt \cdot \int_{-\infty}^{\infty} \left|\mu'(t)\right|^2 dt \tag{1.25}$$

Applying the Schwarz inequality:

$$\int_a^b |x(t)|^2\, dt \cdot \int_a^b |y(t)|^2\, dt \geq \left|\int_a^b x(t)y(t)dt\right|^2 \tag{1.26}$$

for $x(t) = t \cdot |\mu(t)|$ and $y(t) = |\mu(t)|'$, we obtain

$$\delta^2\beta^2 \geq \left(\frac{2\pi}{2E}\right)^2 \left| \int_{-\infty}^{\infty} t\,|\mu(t)| \cdot |\mu(t)|'\,dt \right|^2 \tag{1.27}$$

The latter integral is given by:

$$\int_{-\infty}^{\infty} t\,|\mu(t)| \cdot |\mu(t)|'\,dt = \int_{-\infty}^{\infty} t d\left[\frac{|\mu(t)|^2}{2}\right] = -\int_{-\infty}^{\infty} \frac{[\mu(t)]^2}{2} dt = -E \tag{1.28}$$

Thus, (1.27) yields:

$$\delta \cdot \beta \geq \pi \tag{1.29}$$

This is the *uncertainty principle*, that states that the *time-bandwidth* (TB) product of a signal has a lower limit.

1.6 The time-bandwidth product (TB)

The equal sign in the Schwarz inequality occurs for $y(t) = k\,x(t)$, i.e. when $\mu'(t) = k\,t\,\mu(t)$. The solution of this differential equation gives $\mu(t) = -e^{-kt^2/2}$. Therefore, **the signal with the lowest time-bandwidth product is a single-carrier pulse with a Gaussian envelope.** The important point to be mentioned here is that any modulation will increase the time-bandwidth product. Any waveform with a time-bandwidth product larger compared to unity is referred to as a *pulse compression* waveform. Therefore, the difference between a pulse compression waveform and a single-carrier pulse is the time-bandwidth product. Pulse compression, modulation and high time-bandwidth product all refer to the same property of a signal.

What was illustrated in Fig. 1.1 and Fig. 1.2, was that modulation can increase the signal bandwidth for the same signal duration. The increase in the time-bandwidth product of a signal can serve the need for increasing either the duration or the bandwidth, or both.

In an imaging system, a practical requirement is the utilization of the available system bandwidth . In ultrasound imaging, the system bandwidth is determined by the ultrasound transducer. In a conventional system, this requirement is met by simply transmitting a short pulse. If more bandwidth is available, the single-carrier pulse should be even shorter. This is obviously an inefficient way of using the available bandwidth. A pulse compression waveform can provide the necessary bandwidth without reducing the pulse duration. A short pulse of duration T at a carrier frequency f_0 is a broadband signal containing all frequencies in a bandwidth $B = 1/T$ around f_0. Therefore, the time-bandwidth product of such an unmodulated pulse is in the order of unity, depending on

the definition of the bandwidth. In fact, this is the smallest time-bandwidth product for all signals with the same envelope, and is due to the precise phase relationship of the individual frequency components. If this precise phase relationship is altered, a longer signal will result with the same bandwidth as the pulse. Any non-linear phase function will increase the signal time-bandwidth product. With a proper rearrangement of the phases, the time-bandwidth product of the new signal can be much larger than one.

The last graph of Fig. 1.2 shows that a correlation filter applied on a pulse compression waveform readjusts the phases restoring the short pulse, or more precisely reduces the time-bandwidth product back to an order of one. Additionally, correlation filters are optimal from a detection point of view in the presence of noise. Therefore pulse compression waveforms processed with a correlation-based filter combine the advantages of an optimal system in terms of high axial resolution and signal detection in noise. This is the topic of the following chapter.

CHAPTER

TWO

Pulse compression and the ambiguity function

One of the most important signal processing developments has been that of pulse compression that makes it possible to convert a long frequency-modulated transmitted signal into a much shorter pulse of greater peak power. Pulse compression is based on matched filter theory and allows more efficient use of available bandwidth and transmitted power. This chapter describes the mechanism and basic properties of pulse compression with emphasis on potential application of the technique in ultrasound imaging.

2.1 Filtering using complex notation

A filter is characterized by its transfer function $H(f)$. Since an impulse $\delta(t)$ has a frequency spectrum with an amplitude of one and zero phase for all frequencies, $H(f)$ is simply the impulse response of the filter in the frequency domain. The filter output $G(f)$ for an input signal with spectrum $S(f)$ will be:

$$G(f) = S(f) \cdot H(f). \tag{2.1}$$

The filter output in the time domain is:

$$g(\tau) = \int_{-\infty}^{\infty} S(f) \cdot H(f) e^{j2\pi f\tau} df. \tag{2.2}$$

Using Fourier transform theory, $g(\tau)$ can be readily expressed as a function of the filter impulse response $h(t)$, yielding the fundamental convolution theorem of signal analysis:

$$g(\tau) = \int_{-\infty}^{\infty} h(t) \cdot s(\tau - t) dt = \int_{-\infty}^{\infty} s(t) \cdot h(\tau - t) dt = \mathcal{F}^{-1}\left[S(f)H(f)\right]. \tag{2.3}$$

When the analytic signal in (1.18) $\psi(\tau) = s(\tau) + j\hat{s}(\tau)$ (where $\hat{s}(\tau)$ denotes the Hilbert transform of s(t)) passes through the filter, it can be easily shown that the output is the sum of the real and imaginary responses, i.e.:

$$g(\tau) + j\hat{g}(\tau) = \int_{-\infty}^{\infty} h(t) \cdot s(\tau - t)dt + j \int_{-\infty}^{\infty} h(t) \cdot \hat{s}(\tau - t)dt = \int_{-\infty}^{\infty} h(t) \cdot \psi(\tau - t)dt \tag{2.4}$$

Since the analytic signal contains only positive frequencies, there is no need to use a filter with a real impulse response $h(t)$ that passes both positive and negative frequencies. The complex notation can, thus, be used in a similar manner as in the analysis in Chapter 1, by using a complex filter with impulse response $\eta(t)$ and a one-sided transfer function H(f):

$$\eta(t) = h(t) + j\hat{h}(t) = \int_{0}^{\infty} 2H(f)e^{j2\pi ft}df \quad , \quad \mathrm{H}(f) = \begin{cases} 2H(f) & f \geq 0 \\ 0 & f < 0 \end{cases} \tag{2.5}$$

The complex filter response is then:

$$\gamma(\tau) = \int_{0}^{\infty} \psi(t)\eta(\tau - t)dt = \int_{0}^{\infty} \Psi(f)\mathrm{H}(f)e^{j2\pi f\tau}df \tag{2.6}$$

The advantage of the complex notation is that the envelope of the filter output for the real signal can be written in a compact form as:

$$Env\{g(\tau)\} = \tfrac{1}{2}|\gamma(\tau)| = \tfrac{1}{2}\left|\int_{-\infty}^{\infty} \psi(t)\eta(t - \tau)dt\right| \tag{2.7}$$

2.2 The matched filter

The theory of matched filter was developed after the Second World War in an effort to optimize target detection in radar systems. The matched filter concept is the solution to the problem of finding a linear time-invariant filter, which maximizes the signal-to-noise ratio (SNR) (the peak voltage to noise power) of the receiver output in the presence of white gaussian noise. However, the matched filter is also the ideal receiver from a statistical decision theory point of view (using the likelihood criterion), as well as from an information theory point of view (using the inverse probability criterion). It turns out that Neyman-Pearson and Bayes criteria for optimizing the probability of detection maximize the SNR at the output of the receiver [24]. Thus, the matched filter results in a receiver which maximizes the SNR and at the same time minimizes the probability of error in decision and parameter estimation. Whether such criteria yield the ”optimal” receiver for imaging will be discussed in a following section.

The SNR is defined as the ratio of the peak signal power to the rms noise power:

$$SNR = \frac{\text{peak instantaneous output signal power}}{\text{output noise power}} \tag{2.8}$$

The signal entering the filter is assumed to be corrupted by stationary noise with a uniform spectrum (white noise), i.e. a real-noise waveform $n(t)$ with constant spectral density $N_0/2$. This refers to a two-sided spectrum with both positive and negative frequencies. The matched filter is the optimal filter that maximizes the ratio in (2.8), i.e. the peak power of the filter output signal relative to the noise power. For the derivation of the impulse response of the matched filter using the aforementioned optimization criteria, the reader is referred to Chapter 2 in [25].

The matched filter has an impulse response equal to the input waveform with reversed time axis, except for a gain factor k and a translation in time τ_d for physical realization:

$$h(t) = k \cdot s(\tau_d - t) \quad , \quad \eta(t) = k \cdot y^*(\tau_d - t) \tag{2.9}$$

The transfer function of the matched filter is the complex conjugate of the signal spectrum:

$$\mathrm{H}(f) = k e^{-j2\pi f \tau_d} \cdot \Psi^*(f). \tag{2.10}$$

Therefore, the maximization of the SNR is achieved by the matched filter through the removal of any nonlinear phase modulation and weighting of the received signal spectrum according to the spectral components of the transmitted signal. The matched filter minimizes the effect of noise, by suppressing the noise outside the frequency band of the input signal, essentially making the noise colored.

Substituting the filter transfer function from (2.10) in (2.6), we obtain the matched filter output:

$$\gamma(\tau) = k \int_0^\infty |\Psi(f)|^2 e^{j2\pi f(\tau - \tau_d)} df = R_{\psi\psi}(\tau - \tau_d), \tag{2.11}$$

where (1.16) has been used in the last part of the equation. Thus, if the input signal to the filter was the same as the excitation signal (i.e. if the input noise is negligible), the matched filter response would be mathematically equivalent to the auto-correlation of the transmitted signal, shifted by τ_d.

The maximum SNR occurs at $\tau = \tau_d$ and is equal to:

$$SNR_{\max} = \frac{2 \times \left(\text{ received signal energy } \right)}{\text{noise spectral density } (W/Hz)} = \frac{2E}{N_0} \tag{2.12}$$

The maximum probability of detection is therefore independent of pulse bandwidth and modulation, and depends only on the transmitted energy and the noise power density.

Let the received signal (the input to the matched filter) have an average input power S over the pulse duration T. The average noise power within the input signal bandwidth B is BN_0. For narrowband signals, the instantaneous peak power at any time instance is approximately twice the average power at the same time instance. Then the *matched filter processing gain* or *gain in signal-to-noise-ratio* (GSNR) is given by [26]:

$$GSNR = \frac{SNR_{out}}{SNR_{in}} = \frac{SNR_{\max,out}/2}{SNR_{in}} = \frac{\left(2E/N_0\right)/2}{S/(BN_0)} = \frac{ST/N_0}{S/(BN_0)} = TB. \tag{2.13}$$

The improvement in signal-to-noise ratio will be equal to the time-bandwidth product of the transmitted waveform.

From noise theory, the variance of the output signal phase due to noise is equal to the variance (or average power) of the output noise [23]:

$$\sigma_{\phi}^2 = \frac{1}{E^2}\sigma_N^2 \tag{2.14}$$

The noise power density at the filter output is $(N_0/2)\cdot|\mathrm{H}(f)|^2$ and by substituting H(f) from (2.10) we find the output noise power as:

$$\sigma_N^2 = \int_0^{\infty} \frac{N_0}{4}|\Psi(f)|^2\,df = \frac{EN_0}{2}. \tag{2.15}$$

Substituting (2.15) in (2.14) and converting the phase error into a corresponding delay error, it can be found [1] that the standard deviation σ_τ of a delay measurement is:

$$\sigma_\tau = \frac{1}{\beta\sqrt{2E/N_0}}. \tag{2.16}$$

The accuracy improves with signal-to-noise ratio and this is another important advantage of matched filtering, usually not stressed in the literature. Delay precision improves as the width of the filter response mainlobe decreases, and (2.16) states that this width is inversely proportional to the signal bandwidth. Rihaczek [1] has reported an approximate equation for the half-power width τ_b of the main matched-filter response peak:

$$\tau_b = \frac{\pi/2}{\beta} \tag{2.17}$$

The latter equation can be easily verified heuristically by taking the inverse Fourier transform of $|H(f)|^2$ and then finding its half-power width. For an unmodulated pulse, this is the half-power width of the triangle function shown in Fig. 1.1 and for a rectangular spectrum is the half-power width of the approximate $(\sin x)/x$ output shown in Fig. 1.2.

Thus, the matched filter response has a resolution on the order of the inverse bandwidth. The importance of (2.17) is that in a pulse compression system with modulated signals, range resolution can be set independently of signal duration. Pulse duration can be first adjusted according to power requirements and achievable resolution is only limited by the available system bandwidth. Note, however, the fundamental resolution limitation given by the uncertainty principle.

2.3 Generalized matched filter

The derivation of matched filtering can be extended to non-white or colored noise, i.e. noise that its power density $N(f)$ is a function of frequency. In this case, the filter that maximizes the SNR

at its output has a transfer function [25]:

$$\mathrm{H}(f) = ke^{-j2\pi f\tau_d} \cdot \frac{\Psi^*(f)}{N(f)} \tag{2.18}$$

Assuming that the noise power density $N(f)$ can be factored such as $N(f) = N_m(f) \cdot N_m^*(f) = |N_m(f)|^2$, the generalized matched filter is equivalent to a cascade of two filters, with transfer functions $\frac{1}{N_m(f)}$ and $\frac{\Psi^*(f)}{N_m^*(f)}$ (Fig. 2.1). When the noise passes through the first filter (which is sometimes referred to as the *whitening filter*), its power density becomes

$$N(f) \cdot \left| \frac{1}{N_m(f)} \right|^2 = 1,$$

i.e. the noise at the first filter output is white. At the same time, the signal spectrum is altered by the whitening filter. The second filter is the conventional matched filter, which is matched to the distorted signal in the presence of white noise. Thus, optimal detection in non-white noise is achieved by first whitening the noise and then applying a conventional matched filter on the distorted signal.

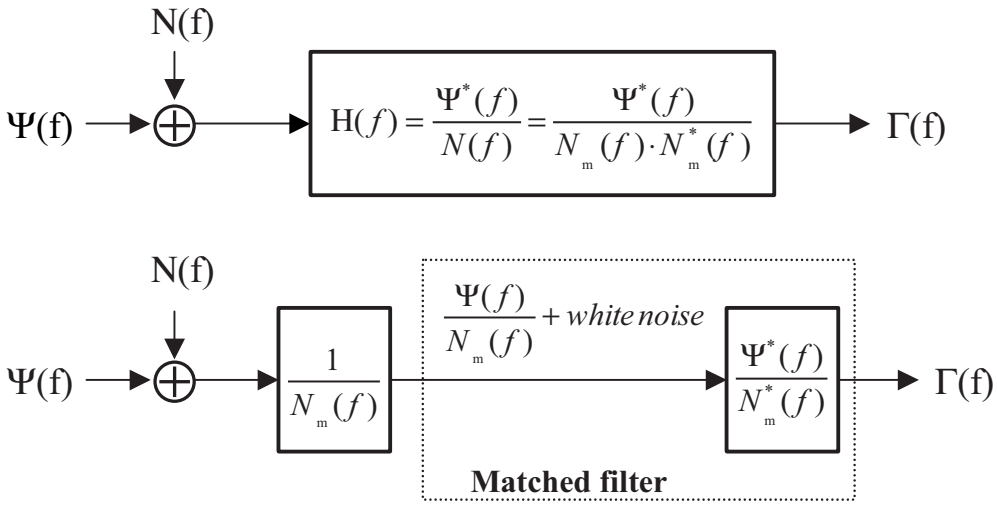

Figure 2.1: Generalized matched filter diagram. In the presence of colored noise, the optimal filter effectively consists of a noise-whitening filter in series with a conventional matched filter.

2.4 Matched filter receiver in ultrasound imaging

2.4.1 Attenuation

Let us assume that a coded waveform $\psi(t)$ is transmitted along direction r into a scattering medium. Let the medium at a given depth r be characterized by its impulse response, a scattering (or reflectivity) function $\gamma(t,r)$. For instance, the impulse response of a single broadband (point) scatterer located at depth r_i will be a delta function:

$$\gamma(t,r_i) = k(r_i)\cdot\delta(t-\frac{r_i}{c}), \tag{2.19}$$

where k is generally a function of depth incorporating acoustic diffraction and attenuation. Assuming that the impulse responses are short, the medium's scattering function $\gamma(t)$ will be the sum of the impulse responses for all depths. The returned signal will then be the convolution of the scattering function with the excitation signal:

$$r(t) = \psi(t) \underset{t}{*} \gamma(t). \tag{2.20}$$

Substituting (2.19) in (2.20), the returned signal from a single scatterer is simply a scaled time-shifted version of the transmitted signal:

$$r(t) = k\cdot\psi(t-\tau_0) = k\cdot\mu(t-\tau_0)\cdot e^{j2\pi f_0(t-\tau_0)} \tag{2.21}$$

where $\tau_0 = r_i/c$ is the time instant after the start of transmission ($t=0$), at which the signal is being received. The factor k is insignificant for the analysis and will be omitted. The actual imaging situation involves an additional convolution of the transmitted waveform with the transducer impulse response. For a transmitted linear FM signal with a rectangular spectrum amplitude, the shape of the propagating pulse spectrum amplitude is effectively that of the transducer. The effect of this weighting will be discussed in details in a following chapter. For the purpose of this discussion, let us assume that (2.21) holds as is, and the received signal is merely a shifted version of the excitation signal.

When the interrogated medium is the human body, there is a frequency-dependent attenuation of the acoustic wave. In this case, the scattering function will be a different function of time for each depth r_i. Several models have been suggested in the literature for the attenuation in tissues [27]. Gurumuthy et.al. [28] have taken into account both the amplitude and the phase of the attenuation, suggesting a minimum-phase impulse response. In the frequency domain, the suggested transfer function $\Gamma(f,r)$ is:

$$\begin{aligned}\Gamma(f,r) &= \exp(-\alpha r)\cdot\exp(-\beta(f-f_0)r) \\ &\times \exp\left(-j2\pi f(\frac{1}{c}+\tau_m\frac{\beta}{\pi^2})r\right)\cdot\exp\left(j\frac{2f}{\pi}\beta r\ln(2\pi f)\right),\end{aligned} \tag{2.22}$$

where α and β are the frequency-independent and frequency-dependent terms of attenuation respectively, expressed in Np/[MHz·cm], and τ_m is the minimum-phase delay factor, with a suggested value of 20 in order to fit the dispersion found in tissue.

In order to find the returned signal from a given depth r_i in an attenuating medium, the inverse Fourier transform of (2.22) has to be convolved with the excitation signal in accordance with (2.20).

The effect of the frequency dependence of the attenuation is a larger attenuation of the high frequencies of the transmitted signal compared to the lower, when this signal propagates through the tissue. That will decrease the upper part of the bandwidth, causing effectively a decrease in the mean frequency. For a simpler model of attenuation and a Gaussian-shaped pulse, Jensen [2] showed that the mean frequency of the propagating pulse decreases linearly with depth:

$$f_{mean} = f_0 - (\beta B_r^2 f_0^2) z, \tag{2.23}$$

where f_0 and B_r are the center frequency and the relative bandwidth of the transmitted pulse respectively, and z is depth in the tissue. A significant downshift can be seen even for moderate tissue depths. For a transducer with a center frequency of 4 MHz and a 65% fractional bandwidth, there is a downshift of 800 KHz at a depth of 16 cm. For moving targets, there is an additional Doppler shift, which is, however, two orders of magnitude less than the frequency downshift due to attenuation.

Assuming that attenuation does not distort the complex envelope of the modulated signal, and it only causes a frequency downshift f_d, a simplified model for the returned signal in an attenuating medium will yield

$$\begin{aligned} r(t) &\simeq \mu(t-\tau_0)\cdot e^{j2\pi[(f_0-f_d)(t-\tau_0)]} \\ &\simeq \mu(t-\tau_0)\cdot e^{j2\pi[f_0(t-\tau_0)-f_d(t-\tau_0)]}. \end{aligned} \tag{2.24}$$

The latter equation is based on the narrowband approximation, i.e. the frequency shift is constant along the entire pulse spectrum.

2.4.2 Matched filter response

Matched filtering of the received signal given in (2.24) requires *a priori* knowledge of the unknown parameters τ_0 and f_d. Practically there is no need to match to the target delay τ_0, since a change in delay merely changes the time at which the output occurs.

In the general case, the returned signal is matched to the signal $r_F(t) = \mu(t-\tau_m)\cdot e^{j2\pi(f_0-f_m)(t-\tau_m)}$ with parameters τ_m and f_m. This corresponds to a filter with an impulse response

$$h_{\tau_m, f_m}(t) = r_F^*(T-t+\tau_m). \tag{2.25}$$

Using the fact that $h_{\tau_m,f_m}(\tau-t)=r_F^*(t-\tau)$, and ignoring the time constant T we get

$$h_{\tau_m,f_m}(\tau-t)=\mu^*(t-\tau-\tau_m)\cdot\exp\{j2\pi[f_0(-t+\tau+\tau_m)+f_m(t-\tau-\tau_m)]\} \tag{2.26}$$

Substituting the returned signal given in (2.24) and the filter impulse response from (2.26) into (2.6), we get the receiver output:

$$\begin{aligned} x(\tau) &= \int_{-\infty}^{\infty}\mu(t)\cdot\mu^*(t-\tau+\tau_0-\tau_m)\cdot\exp\{j2\pi[(f_0-f_m)(\tau-\tau_0+\tau_m)+(f_m-f_d)t]\}dt \\ &= e^{j2\pi(f_0-f_m)(\tau-\tau_0+\tau_m)}\cdot\chi(\tau-\tau_0+\tau_m,f_m-f_d) \end{aligned} \tag{2.27}$$

where $\chi(\tau,f_d)$ is defined as the *ambiguity function* , and is given by

$$\chi(\tau,f_d)=\int_{-\infty}^{\infty}\mu(t)\cdot\mu^*(t-\tau)\cdot e^{j2\pi f_d t}dt=\int_{-\infty}^{\infty}\psi(t)\cdot\psi^*(t-\tau)\cdot e^{j2\pi f_d t}dt. \tag{2.28}$$

If the delay is translated by $\tau_0-\tau_m$, so that the maximum occurs at $\tau=0$ rather than at $\tau=\tau_0-\tau_m$, and the frequency is translated by f_m, so that matching occurs at $f_d=0$ rather than at $f_d=f_m$, the ambiguity function is the matched filter response. The value of the ambiguity function at point (τ,f_d) away from the origin $\tau=f_d=0$ shows the response of a filter which is mismatched by τ and f_d relatively to the input signal.

2.5 The ambiguity function and its properties

The ambiguity function (AF) given in (2.28) is the complete 2-D matched-filter response in scatterer range and frequency downshift. An image of a number of scatterers will be the superposition of weighted and time-shifted ambiguity functions. Obviously, the ideal ambiguity function is a two-dimensional delta function located at the origin. However, from a mathematical point of view, there are some strong restrictions for a two-dimensional function to be an ambiguity function, and as it will shown in this section, a delta function is unrealizable. The objective of waveform design is therefore the control of the shape of the ambiguity function for a given imaging medium. This section gives the basic properties-constrains of the ambiguity function. For a more detailed discussion, the reader is referred to [25].

In the absence of attenuation (or generally when $f_m-f_d=0$), the matched filter output is the autocorrelation function of the waveform. The maximum of the ambiguity function occurs at the origin:

$$\chi(0,0)=\int_{-\infty}^{\infty}|\psi(t)|^2dt=\int_{-\infty}^{\infty}|\mu(t)|^2dt=2E \tag{2.29}$$

The squared modulus of the ambiguity function is referred to as the *ambiguity surface* (AS) [1]. The ambiguity surface is symmetrical around the origin, i.e. $|\chi(-\tau,-f_d)| = |\chi(\tau,f_d)|$. The volume under the ambiguity surface is given by [1, 25]:

$$V_{amb} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |\chi(\tau,f_d)|^2 \, d\tau d\, f_d = \left| \int_{-\infty}^{\infty} |\mu(t)|^2 \, dt \right|^2 = |\chi(0,0)|^2 = 4E^2 \tag{2.30}$$

This equation states that the total volume under the AS is only a function of the signal's energy and therefore independent of modulation. It also states that the volume under the surface is equal to the maximum squared. If a waveform is chosen whose AF has a narrow central spike, the bulk of the fixed volume under the ambiguity surface must appear as a pedestal on the $\tau - f_d$ plane around the spike. This AF is referred to as the "thumbtack" function in the radar literature. This sets a fundamental limit on achievable resolution.

In addition to the property of the total volume invariance, the AF is constrained on the distribution of this volume in τ and f_d. The ambiguity surface is its own 2-D Fourier transform. The integrated volume distribution in frequency shift is completely determined by the value of the ambiguity function on the delay axis. This is based on the following property of the ambiguity function [1]:

$$\int_{-\infty}^{\infty} |\chi(\tau,f_d)|^2 \, d\tau = \int_{-\infty}^{\infty} |\chi(\tau,0)|^2 \, e^{-j2\pi f_d \tau} d\tau \tag{2.31}$$

More precisely, the velocity distribution of the ambiguity volume is the Fourier transform of the squared envelope of the ambiguity function on the delay axis, i.e. the squared envelope of the autocorrelation function of the transmitted signal. The dual integral of (2.30) states that the integrated volume distribution in delay is given as the Fourier transform of the ambiguity surface on the frequency shift axis, which depends only on the autocorrelation function on this axis. Therefore this distribution is independent of any phase modulation for a given envelope $\mu(t)$. What modulation does, is simply to redistribute the fixed volume in each delay slot, but no volume can be moved from one slot to another.

These transform relations imply that if the central peak is narrowed along one axis, the volume must spread out in the other axis.

2.6 Classification of pulse compression waveforms

2.6.1 Classification based on their coding

The time-bandwidth product of a waveform can be increased by modulating either the amplitude or the phase. Amplitude modulation is suboptimal in terms of signal energy, and results in signals

that pose linearity requirements on the amplifiers over the entire dynamic range of the echoes, and will not be considered further. The phase-modulated signals can be categorized in two main groups, based on the kind of phase modulation that is applied: phase modulation can be either continuous or discrete. This gives the main classification of pulse compression waveforms in analog (or frequency-modulated (FM)) signals and discrete (usually phase-modulated) codes. A third way to increase the time-bandwidth product is the pulse trains.

FM waveforms are the modulated signals, in which the phase of the waveform changes continuously with time, so that the definition of the instantaneous frequency given in (1.3) is meaningful. The phase modulation of FM signals is a continuous non-linear function of time. The simplest and most important of all FM waveforms is the linear FM with a quadratic-phase modulation (and therefore linear instantaneous frequency). Other commonly used non-linear FM signals is the cubic FM and the sinusoidal frequency modulated signal (SFM). The family of FM waveforms will be discussed in Chapter 3.

Discrete coded waveforms are signals, in which the phase changes in discrete steps, usually every one or more cycles of a cw signal. In spread spectrum application, this coding is called phase shift coding or keying (PSK). Frequency shift keying (FSK) yields FM discrete waveforms where the frequency changes in discrete steps. Shifting of the phase in the frequency domain rather than in the time domain is also possible, but results in waveforms with non-constant amplitude and practically is not used. Binary or phase-reversal codes are referred to as the phase-modulated codes where the phase takes only two values: 0° and 180°. Polyphase codes can have more than two-value phase encoding. Discrete coded waveforms will be discussed in Chapter 5.

Pulse trains consist of repetition of a single pulse. If the same waveform is used from pulse to pulse, these waveforms are called uniform pulse trains. Frequency or phase coding from pulse to pulse is also possible or repetition of a signal which has itself a high time-bandwidth product. Coding of the interpulse duration (called staggering) offers another degree of coding. Apparently waveforms with extreme complexity are possible by combining several of the aforementioned techniques.

Repetition of pulses can also be contiguous, although such signals can fall into the category of CW or phase-modulated signals. Therefore, for this analysis, pulse trains will be considered the signals which are a repetition of pulses with a duty cycle less than one. However, a low duty-cycle implies either long transmitted signals or inefficient transmitted energy and their usefulness in ultrasound imaging may be limited. The utility of pulse trains can be in processing echoes from several emissions with a different coding from emission to emission with a single matched filter. Pulse trains and their potential in ultrasound imaging will be discussed in detail in Chapter 10.

2.6.2 Classification based on their ambiguity function

The previous classification of pulse compression waveforms is based on the kind of modulation that is applied on the transmitted waveform. However, the relevant quantity for imaging is the corresponding ambiguity function of the coded waveform. Signals can be categorized [29] according to their ambiguity function into basically four groups:

- single-carrier pulses
- waveforms with thumbtack ambiguity function
- waveforms with a ridge ambiguity function
- waveforms with AF of discrete ambiguities and clear areas

Constant-carrier pulses

Fig. 2.2 shows the ambiguity function of a short and a long pulse. The short pulse is 4 cycles of a 4 MHz sinusoidal signal. The axial resolution (on the time axis) is on the order of the inverse bandwidth, and the frequency resolution is on the order of the inverse duration of the pulse.

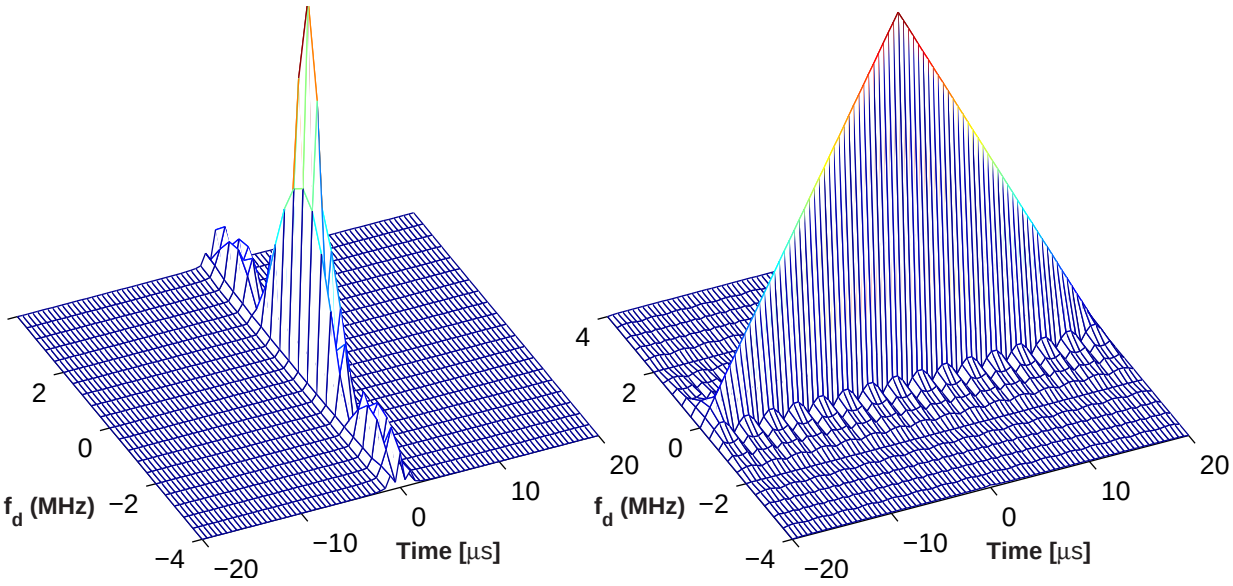

Figure 2.2: The ambiguity functions of unmodulated pulses have a triangular shape on the time axis and a *sinc* shape on the frequency axis.

Waveforms with thumbtack ambiguity function

If it is desired to measure simultaneously from a single measurement the range and Doppler of a point scatterer in a non-attenuating medium, the ideal ambiguity function would be that of a narrow spike located at the center. The range resolution of such spike will be proportional to the inverse bandwidth, while the Doppler resolution will be proportional to the inverse of the duration of the waveform. From the uncertainty principle, it can be seen that the spike cannot be made arbitrarily narrow. The narrower it is on one axis, the broader will be on the other. Additionally, the volume constraint of the ambiguity surface discussed on the last section, will result in a high pedestal where the bulk of the ambiguity surface volume will be concentrated. The volume distribution on one axis is the Fourier transform of the ambiguity surface on the other axis. For a waveform with a spike width on the delay axis of $1/\beta$, most of the volume will be spread on a frequency interval β. In an analogous manner, if the spike width on the frequency axis is $1/\delta$, the pedestal will be spread on the delay axis over δ. The pedestal level is inversely proportional to the time-bandwidth product of the waveform. When the time-bandwidth increases, the pedestal level gets lower, however, it is extended in a longer range. The total volume within the pedestal is $\delta\beta \cdot (1/\delta\beta) = 1$, while the volume in the central spike is $1/\delta\beta$. Such ambiguity function is known in the radar literature as the ”thumbtack” function , because of its shape. A sketch of the thumbtack ambiguity function is illustrated in Fig. 2.3.

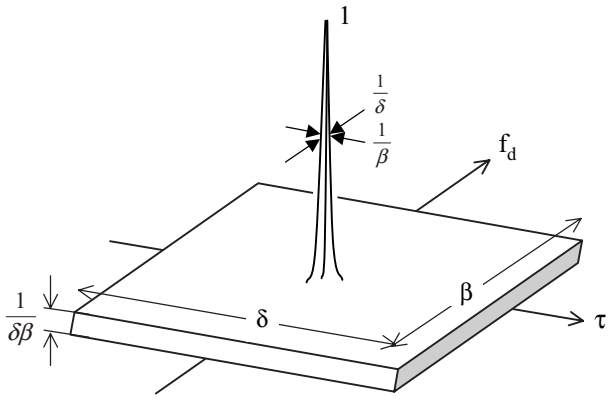

Figure 2.3: Sketch of the ideal thumbtack ambiguity function [1].

The thumbtack ambiguity function offers good resolution both in range and frequency shift, with two trade-offs: i) a high pedestal and ii) that when the frequency shift mismatch exceeds $1/\delta$, effectively no compression occurs. Waveforms whose ambiguity function resembles the thumbtack function are generally those in which some kind of random coding is involved, such as the pseudo-random codes. Fig. 2.4 shows the ambiguity function of an m-sequence (maximum-length

sequence) of length 64.

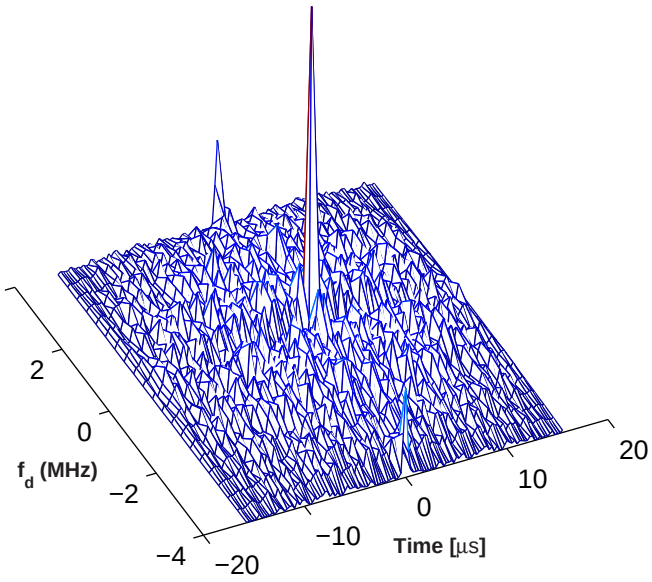

Figure 2.4: The ambiguity of an m-sequence of length 64.

Waveforms with a ridge ambiguity function

These waveforms have the property that pulse compression still occurs in the presence of a mismatch. The effect of a frequency shift is only a time translation of the compressed pulse. This can happen only if the matched filter has a group delay which is proportional to frequency, i.e. a quadratic phase function. Waveforms with this property are the linear FM, discrete phase-coded waveforms (such as the Frank polyphase code) where the carrier frequency is stepped in a linear manner, as well as coherent pulse trains with linear frequency shifting from pulse to pulse. The time translation of the matched filter output is proportional to the frequency shift mismatch, and the ambiguity function consists of a long ridge with a slope β/δ (Fig. 2.5).

Waveforms with AF of ambiguous spikes and clear areas

The ambiguity function of a pulse train, usually referred to as the ”bed-of-nails” ambiguity function is shown in Fig.2.6. It consists of an array of regularly spaced spikes with the spacing on the delay axis equal to the pulse repetition period and the spacing on the frequency shift axis equal to the repetition frequency. The resolution of each of the spikes is the same as in the thumbtack function,

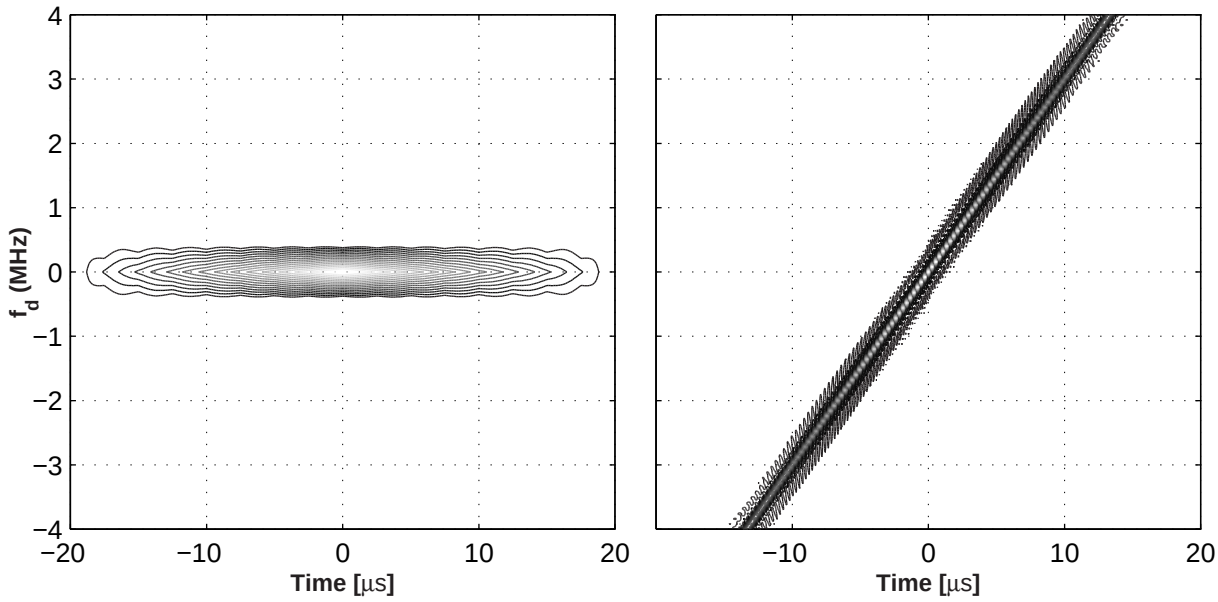

Figure 2.5: Contour plots of the ambiguity functions for a single-carrier pulse (left) and a linear FM (right) with the same duration. The left graph is the contour plot of Fig.2.2. The presence of the linear frequency modulation shears the ridge away from the delay axis. The slope of the ridge is β/δ.

i.e. $1/\beta$ and $1/\delta$. However, the self-noise energy is not spread on a uniform pedestal, but is concentrated in isolated regions.

2.7 Resolution in a matched filter receiver

2.7.1 Resolution and resolvability

The problem of imaging with modulated signals is two-fold: detection and resolution. Detection is optimized by maximizing the SNR. This is achieved by the matched filter receiver.

The second is that the filter output should have good axial resolution. For example, a long non-modulated pulse processed with its matched filter gives the triangle output shown in 1.1. This system although it maximizes the SNR at the filter output, is inappropriate in terms of resolution.

Resolution is usually defined as the ability of an imaging system to separate two closely spaced scatterers based on Raleigh or other criteria. This is determined only by the mainlobe of the autocorrelation function. For a rough estimation of resolution, a usual index is the half-width of the mainlobe. As described previously, for a matched filter receiver in the presence of attenuation, resolution at a given depth is determined by the mainlobe of a cut of the ambiguity function parallel

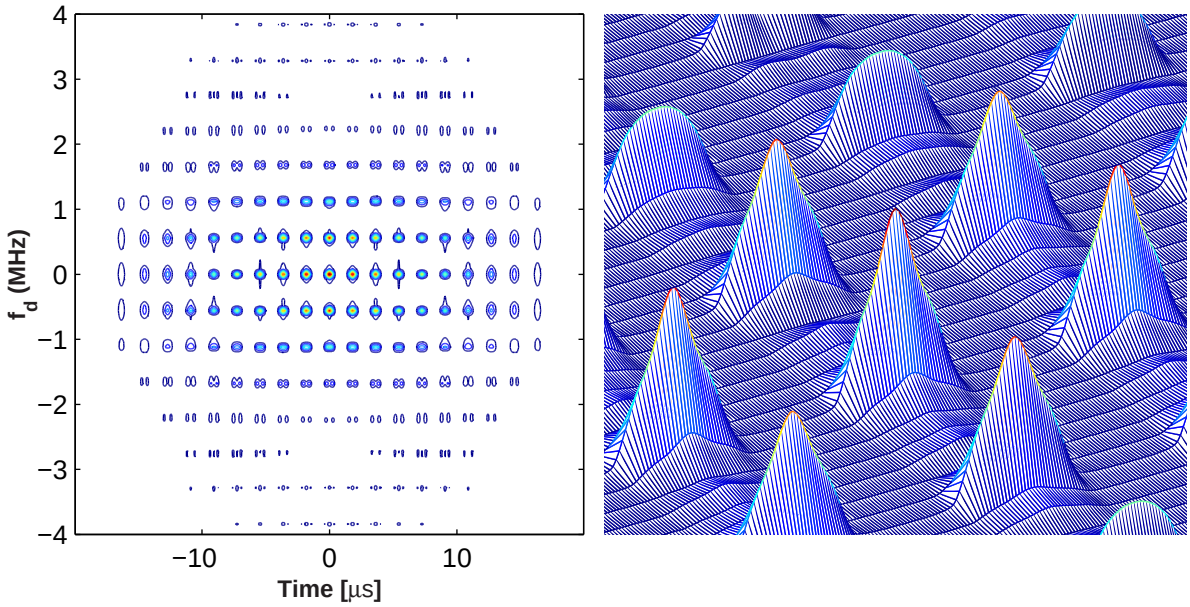

Figure 2.6: Contour plot of the ambiguity function of a pulse train (left) and detail of the central part (right). The pulse train consists of 11 2-cycle pulses with a duty cycle of 0.275.

to the delay axis at the frequency shift in this depth.

However, when modulated signals are used, sidelobes are introduced. This sidelobe behaviour of the point-spread function is referred to in the radar literature as the *waveform self-clutter* or *self-noise*, as well as *resolvability* or *discrimination* [30]. If the sidelobes are high, a weak scatterer being imaged in the presence of a much stronger scatterer may not be resolved or distinguished, because it is hidden in the sidelobe region of the ambiguity function. Frequently used indices are the peak sidelobe level (PSL) and the integrated sidelobe level (ISL). PSL is simply the ratio between the highest sidelobe and the mainlobe. ISL is the ratio of the total sidelobe energy to main peak energy expressed in decibels. ISL describes the leakage of energy from bright into dark areas in the axial direction of an image [31].

Therefore resolution in a matched filter system is a combination of two factors: close-target separability (associated with the mainlobe response) and self-clutter (associated with the sidelobes in the $\tau - f_d$ plane).

The mainlobe width is proportional to the inverse of the bandwidth, and thus, axial resolution can be improved by broadening either the spectrum of the matched filter or that of the transmitted waveform. The former case will be accompanied by enhancement of noise, and therefore it is preferred to transmit a waveform with as wide a spectrum as possible. In ultrasound, system bandwidth is limited by the frequency response of the transducer, which limits the degree which the mainlobe can be narrowed to. From the discussion in Section 1.6, it can be generally concluded

that, as long as the entire available bandwidth is utilized, axial resolution with coded signals will be at the same order as for conventional pulse imaging.

Supression of the so called range sidelobes is a key issue in resolvability in coded imaging. Since the impulse response of the matched filter is the time reversal of the transmitted waveform, the design of the optimal imaging system has as a prerequisite the appropriate choice of the transmitted waveform. Design strategies for sidelobe reduction depend on the used coded waveform and will be extensively discussed in this work: in Ch.4 for the family of FM signals, in Ch.5 for the discrete codes, and in Ch.10 for pulse trains. Assuming matched filtering on receive, the usual approach is to design waveforms whose auto-correlation function has low sidelobes, since the auto-correlation function is the zero-frequency shift case of the ambiguity function. In this chapter, the matched filter was derived as the optimal receiver for detection in the presence of noise. However, the matched filter is not optimized for detection of scatterers in the self-clutter. For example, the matched filter output for zero frequency shift of a linear FM has range sidelobes that are only 13.2 dB lower than the mainlobe. Non-linear FM modulation can give sidelobes down to -35 dB. One of the tasks in coded imaging is, thus, the design of suitable waveforms with good range sidelobe properties.

Apart from the waveform selection, a second possibility in suppressing the range sidelobes is by slightly modifying the matched filter to a mismatched condition. The general principle and the implications of mismatched filtering in resolution will be discussed qualitatively in the following section.

2.8 Mismatched filtering

The transfer function amplitude of the matched filter has the spectrum amplitude of the transmitted waveform. Eq.(1.20) shows that the envelope of the output signal is the inverse Fourier transform of the squared spectrum amplitude. For a linear FM signal with a rectangular spectrum amplitude, the output is approximately a *sinc* function as shown in Fig. 1.2. It is, thus, the sharp edges of a rectangular spectrum that generate the high slowly-decreasing sidelobes of the *sinc* function. If a window function is applied on the transfer function amplitude of the matched filter, these sidelobes can be reduced almost at will. An excellent discussion on windowing functions and their sidelobe properties can be found in [31]. The weighted matched filter is usually referred to as the *mismatched* filter.

There are mainly two drawbacks of mismatched filtering: The first is a widening of the mainlobe. For ultrasound, even when heavy weighting is applied, the widening will not exceed 2 times the width of the unweighted mainlobe; a resolution which is acceptable in most cases. Design guidelines for the trade-off between sidelobe level and mainlobe width will be given in following chapters. The second undesired effect of mismatched filtering is a degradation in signal-to-noise

ratio, which is -for useful ultrasound signals - around 1.5 to 3 dB compared to the matched case. The sidelobes effects are however more deleterious, and a sacrifice in that range in SNR is often a reasonable price to pay.

For small frequency shifts, mismatched filtering is equivalent to the transmission of a weighted spectrum. In ultrasound, such weighting occurs because of the transducer frequency response, resulting in reduced sidelobes without any mismatched filtering. Further signal predistortion for optimal spectral weighting in the special case of FM signals will be discussed in Chapter 4. In the presence of frequency shifts, an amplitude-weighted receiver filter passes an unsymmetrical spectrum, distorting the weighting effect and can introduce high sidelobes. In this case, it is necessary to weight both the transmitted spectrum and the filter spectrum, a design often called *bilateral* weighting.

Although mismatch in the phase function rather than in the amplitude is possible, phase mismatching is a sensitive operation and there is an increased difficulty in controlling the resulting sidelobes. Even when the amplitude is heavily tapered, the mismatch is small compared to the mismatch in the phase [1].

If $\mu_2(t)$ is the modulation function of a mismatched filter to a transmitted signal with modulation function $\mu_1(t)$, the *cross-ambiguity* function can be defined in accordance to the ambiguity function:

$$\chi_{12}(\tau, f_d) = \int_{-\infty}^{\infty} \mu_1(t) \cdot \mu_2^*(t-\tau) \cdot e^{j2\pi f_d t} dt. \tag{2.32}$$

Rihaczek [1] showed that the volume of the cross-ambiguity surface is constant for a given filter if the input signal is varied so that its energy remains unchanged. That implies that mismatched filtering does not release the volume constraints of the cross-ambiguity function. He also showed that the average self-clutter level of a cross-ambiguity function cannot be lower than that of the ambiguity function, i.e. the effective decrease in the filter passband does not solve the self-noise problem on the (τ, f_d) plane. However, mismatching can be very useful in suppressing sidelobes on a plane along the delay axis or shifting part of the volume into regions with less interference, at the expense of a slight increase in the mainlobe width.

It can be summarized that the problem of designing a suitable transmitted waveform and a suitable ”mismatched” filter for coded ultrasound imaging at a given depth is that of a trade-off between the sharpness of the central spike, the spreading of the undesired self-clutter in the ambiguity surface, the SNR degradation due to the mismatch, and the degree of deviation in the filter response for other frequency shifts.

2.9 Optimal filtering in speckle

Speckle

The human tissue can be considered as composed by an ensemble of scatterers, which are spaced much closer than the effective waveform resolution at the matched filter output and thus cannot be resolved. The resulting speckle pattern in an ultrasound image is only a constructive and destructive interference (depending on the phases) of the scattering strengths. The speckle signal is deterministic in the sense that observations are not independent, i.e. the same signal will result for repeated measurements. That implies that averaging cannot reduce the speckle signal as does with noise. Additionaly, the signal-to-thermal noise-ratio can be increased by increasing the transmitted power, but this does not change the signal-to-clutter-ratio, since speckle is a result of a convolution process.

However, the phase of the speckle is random whereas the amplitude is dependent on the signal spectrum amplitude. Speckle noise has been referred to as multiplicative noise [26], as opposed to thermal noise which is additive. For a large number of randomly distributed scatterers, the speckle signal has gaussian character, and if it is uniformly distributed in range, it can be considered as stationary gaussian noise [1]. The difference between thermal noise and clutter is that the power spectral density of speckle is frequency dependent, so that the speckle ”noise” is nonwhite. Assuming independent scatterers, the echoes will have the same frequency spectrum as that of the transmitted signal, and the power density spectrum of speckle N_c will be: [25]:

$$N_c(f) = k_c \left|S(f)\right|^2, \tag{2.33}$$

where k_c is the average scattering strength.

2.9.1 Inverse or equalization filter

The returned signal $r(t)$ is given in (2.20) by the convolution of the transmitted signal $\psi(t)$ with the scattering function $\gamma(t)$. In a matched filter receiver, $r(t)$ is correlated with the transmitted signal:

$$x(t) = \psi^*(t) \underset{t}{*} r(t) = \psi^*(t) \underset{t}{*} \psi(t) \underset{t}{*} \gamma(t), \tag{2.34}$$

where $x(t)$ is the matched filter output. The objective of imaging is the extraction of the scattering function $\gamma(t)$. However, the output of the matched filter is the convolution of $\gamma(t)$ with the autocorrelation function of the transmitted signal. The autocorrelation function is, thus, the point spread function or the imaging kernel of the matched filter imaging system in the axial direction. In cases where SNR is not a problem, the ideal filter should have a transfer function equal to the inverse of

the input spectrum:

$$H(f) = \frac{1}{\Psi(f)} \tag{2.35}$$

In this way, the convolution of the signal with the reflectivity function of the medium is cancelled. In practice, such inverse filter cannot be realized, because it has infinite response outside the signal bandpass, and implementation is based on approximation of the inverse filter only over a limited frequency range (pseudo-inverse filter). By suppressing the passband response, the inverse filter effectively widens the bandwidth and thereby improves the axial resolution, with the expense of amplifying the relative noise power and degrading the SNR.

Noise-free speckle

When the medium consists of speckle with power density spectrum given by (2.33), the generalized matched filter of (2.18) gives the inverse filter. The inverse filter will be the optimal filter when no noise is present. This is of course not a realistic condition, especially when the spectrum amplitude is a weighted function as it is when it passes through the transducer. Even when the noise is low, it will be significant at the edges of the spectrum.

2.9.2 Wiener filter

The Wiener filter is derived as the solution to the optimization criterion of the best mean-square fit to a desired signal over a given time or frequency range, usually with the best fit over the passband of the input signal. For an input signal with spectrum $\Psi(f)$ in the presence of noise with a spectrum $N_f(f)$, the Wiener filter has a transfer function [22, 5]:

$$\mathrm{H} = \frac{\Psi^*}{\Psi\Psi^* + N_f N_f^*}. \tag{2.36}$$

Speckle and noise

If the generalized matched filter is applied to the problem of detection in the presence of speckle, when gaussian receiver noise is added to the filter input, the optimum receiver should have a transfer function:

$$H(f) = \frac{\Psi^*(f)}{N_0/2 + k_c|\Psi(f)|^2} \tag{2.37}$$

A comparison of the last equation with (2.36) shows that in this imaging situation the optimum filter is the Wiener filter. In the absense of clutter, the Wiener filter is simply the matched filter,

while in the absence of noise it is the inverse filter. The transfer function at the outskirts of the spectrum resembles that of the matched filter, while at the passband resembles that of the inverse filter. Therefore, what the filter does is to suppress the bandpass region and effectively to enhance the outskirts of the spectrum. In this way, the clutter power is spread and the peak signal-to average clutter ratio is improved. Close to the edges, where the clutter is not significant, but the noise is, the filter is matched to the signal.

The problem in constructing such a filter is that the mean scattering strengh k_c relatively to the noise level is not known. The performance of the filter depends on the clutter-to-noise ratio which is unknown and maybe variable depending on the interrogated tissue.

2.10 Appropriate compression waveforms and filters for ultrasound imaging

In the literature there is a lack of systematic methodology to devise optimum waveforms. In most cases, the ambiguity function (or the auto-correlation, which is the zero frequency shift case of the ambiguity function) of several waveforms is computed and its properties have been studied using *ad hoc* or trial-and-error approaches. There are only few cases of waveform synthesis based on the properties of the corresponding ambiguity function. [32, 33]. Wolf et. al. [33] have studied the problem of finding waveforms which yield an ambiguity function that is the minimum mean-square estimate of the desired ambiguity function over any specified region on the (τ, f_d) plane.

Once a waveform is chosen that is associated with a suitable ambiguity function, the matched filter receiver assures detection optimality. The volume invariance constraint of the ambiguity function implies that all signals with the same energy are equally good (or bad) in terms of the total ambiguity. The choice of an appropriate waveform depends on the scattering medium.

2.10.1 Detection of isolated scatterers in white noise

Matched filtering has been designed for detection of isolated scatterers in white noise, and is therefore the optimal receiver in this case. In radar systems, the objective of pulse compression is the detection of isolated targets in noise and the estimation of range and velocity of each of the targets with minimum rms error. The task for the radar designer is to design signals and corresponding ambiguity functions that have good resolution properties in both range and Doppler. If the medium consists of only few scatterers, the thumbtack AF would be the appropriate choice.

In ultrasound, although attenuation causes a frequency shift to the returned echoes similar to the Doppler effect, there is no need for retaining good resolution in the frequency shift axis. The reason for that, is that blood velocity cannot be measured anyway in ultrasound from a single mea-

surement, since the Doppler shift is obscured by the frequency shift from the attenuation. Blood velocity is measured from the phase shift between successive pulses [2]. Therefore, the ideal ambiguity function for ultrasound should be as insensitive to frequency shifts as possible, since the frequency shift of the returned signal due to attenuation is not known. That precludes the so called in radars ”Doppler-sensitive waveforms”, that their ambiguity function is an approximation of the thumbtack function, such as PN sequences, Barker codes etc.[1] In ultrasound, a coded waveform should only have good range resolution for a large range of frequency shifts. Frequency shift resolution is inversely proportional to the pulse duration, while range resolution is inversely proportional to the pulse bandwidth.

A good choice for ultrasound imaging would then be a very short pulse. It has good range resolution because of its large bandwidth and is insensitive to frequency shifts due to its short duration. Note that if both range and frequency shift are to be measured, as in radar systems, such a waveform is inappropriate. As it was mentioned before, the problem with the short pulse excitation is its low transmitted energy and low time-bandwidth product. High TB results in high signal-to-noise ratio, which in turn provides the optimal detection and measurement accuracy. As it was shown in the previous chapter, a Gaussian short pulse would be the worst possible choice in terms of SNR, since it has the minimum time-bandwidth product.

The linear FM signal is a good candidate as long as appropriate weighting assures good range resolution and sidelobe suppression. Compression is satisfactory for a large range of frequencies due to the ambiguity ridge and is ideal for uncompensated frequency shifts. The rectangular signal spectrum allows full utilization of the system bandwidth (as opposed to other candidates such as the pulse trains) and good SNR improvement.

If the medium consists of several semi-isolated scatterers, which are placed closer than $1/T$ (where T is the transmitted pulse duration), waveforms with a thumbtack ambiguity function are not well suited, due to the uniform pedestal. For a waveform 16 μs long at 4 MHz with a thumbtack ambiguity function, the pedestal level will be about -36 dB below the peak output. For practical realizable waveforms with non-ideal thumbtack ambiguity, this level will be around -30 dB. Then a 30 dB weaker scatterer in the vicinity of a strong scatterer will be completely masked.

2.10.2 Detection in speckle and white noise

If the problem is detection of signal in the presence of speckle, as is the case of detection of tumors in soft tissues such as liver or breast, the design considerations involve the choice of the appropriate waveforms and the optimization of the filter.

[1] A use of a frequency-shift-sensitive waveform in ultrasound can be a method for estimating ultrasound attenuation. Since such a waveform has good frequency resolution, a bank of matched filters, each tuned at a different frequency can yield the frequency shift due to attenuation for a given depth.

From the discussion of the ambiguity function properties, it is clear that waveforms with a high pedestal will be inappropriate for imaging a dense scatterer environment as in ultrasound. The self-clutter of the thumbtack function will build up degrading the resolution. It can then be preferable to concentrate the volume into a sharp ridge as the linear FM signal does, or to use a pulse train if the strong discrete ambiguities can somehow be resolved.

The optimal filter in this case is the Wiener filter described in Section 2.9.2. Rihaczek [1] has studied the performance loss for the matched and inverse filters over the optimal Wiener filter as a function of the clutter-to-noise ratio. By choosing a realistic weighted spectrum, he calculated the SNR loss for the three filters under the constraint of constant power, and reported a performance loss of 1.4 dB for the inverse filter and only 0.48 dB for the matched filter. This should not be a surprising result, since the phase characteristics of all three filters are similar. Considering the practical difficulty of selecting the clutter-to-noise ratio in the Wiener filter, Rihaczek questioned the usefulness of using an optimal filter instead of the matched filter.

The optimality of the transfer function of the Wiener filter lies on a relative amplification of the spectrum outskirts. This effect can be achieved by widening the bandwidth of the transmitted signal slightly over the bandwidth of the transducer.

Manasse showed that the optimization of the SNR for a fixed signal bandwidth is a signal with rectangular spectrum for the optimal filter [25]. He also showed that weighting causes only small degradation, which for the matched filter is independent of the speckle-to-noise ratio. From all compression waveforms, the one with a rectangular amplitude spectrum is the linear FM signal. Therefore, a linear FM signal processed with a mismatched filter would be an appropriate choice for speckle and receiver noise. These notes indicate that it is preferable to use a proper signal and mismatched filter design than any attempt of using the optimal but speckle-dependent Wiener filter.

In the presence of attenuation, ideally a range-variant matched filter has to be applied. This is a big computational load for the receiver. If a linear FM signal is used, a single matched filter can be used for all depths, with no need to compensate for the frequency shifting caused by attenuation. This reduces significantly the required signal processing.

CHAPTER

THREE

The linear FM signal and other FM waveforms

Almost any waveform with some kind of frequency modulation provides pulse compression, when processed with a matched filter. The simplest case is the linear FM signal, where its instantaneous frequency is a linear function of time. Among all compression waveforms, the linear FM or chirp signal is the one that has been studied more extensively, particularly in the radar literature. One of the reasons for that was the ease of generation in the analog world of the compression radar systems. However the most important reason is its unique symmetry properties and their implications on pulse compression, which will be discussed in details in this chapter. The linear nature of this waveform makes it rather insensitive to frequency shifts as opposed to non-linear chirps and biphase codes; a property which is of great importance in ultrasound imaging.

This chapter is organized as follows: First the spectrum of the linear FM signal is derived analytically and the symmetry properties are discussed. Then, the analytical expressions for the matched filter response and the ambiguity function of the linear FM are given. The properties of compression, resolution and sidelobe levels are then discussed. Techniques for sidelobe reduction are left to presented in the following chapter. Some comments are given on signal-to-noise ratio, its theorectical expected improvement and gain losses resulting from weighting. Finally, non-linear FM signals are considered and discussed briefly. This chapter forms the basis for Chapter 6, where ultrasound imaging with these waveforms will be evaluated.

3.1 The linear FM signal

The general linear FM (or *chirp*) signal can be expressed in complex notation as:

$$\psi(t) = a(t) \cdot \exp\left[j2\pi\left(f_0 t + \tfrac{B}{2T}t^2\right)\right], \quad -\tfrac{T}{2} \le t \le \tfrac{T}{2} \ , \tag{3.1}$$

where f_0 is the center frequency, T is the signal duration and B is the total bandwidth that is swept. Applying (1.3), the instantaneous frequency is given by:

$$f_i = \frac{d\left(f_0 t + \frac{B}{2T}t^2\right)}{dt} = f_0 + \frac{B}{T}t. \tag{3.2}$$

The linear FM signal has a quadratic phase modulation function and, therefore, a linear instantaneous frequency function of time. Although the term "instantaneous frequency" contradicts with Fourier theory, where a signal has to be infinite in time in order to be band-limited, in practical terms f_i indicates the spectral band in which the signal energy is concentrated at the time instant t. The parameter B/T is often denoted as μ, and referred to as the *FM slope* or the *rate* of FM sweep. The signal sweeps linearly the frequencies in the interval $[f_0 - \frac{B}{2}, f_0 + \frac{B}{2}]$.

For digital generation of the FM signal, an alternative expression that has no negative times has the form:

$$\psi(t) = a(t) \cdot \exp\left\{ j2\pi \left[(f_0 - \tfrac{B}{2})t + \tfrac{B}{2T}t^2 \right] \right\} \quad , \quad 0 \le t \le T \ . \tag{3.3}$$

3.2 Spectrum of the linear FM signal

For a linear FM signal with rectangular envelope ($a(t) = 1$), its Fourier transform gives:

$$\begin{aligned} \Psi(f) &= \int_{-T/2}^{T/2} \exp[j2\pi(f_0 t + \frac{1}{2}\underbrace{\left(\frac{B}{T}\right)}_{\mu} t^2)] \cdot \exp(-j2\pi f t)\, dt \\ &= \int_{-T/2}^{T/2} \exp\left\{ j2\pi \left[(f_0 - f)t + \frac{\mu}{2}t^2 \right] \right\} dt \end{aligned} \tag{3.4}$$

Completing the square in the brackets:

$$(f_0 - f)t + \frac{\mu}{2}t^2 = \frac{\mu}{2}\left(t - \frac{f - f_0}{\mu}\right)^2 - \frac{(f - f_0)^2}{2\mu},$$

the integral becomes:

$$\Psi(f) = \exp\left[-j\frac{\pi}{\mu}(f - f_0)^2 \right] \int_{-T/2}^{T/2} \exp\left[j\frac{\pi}{2}2\mu\left(t - \frac{f - f_0}{\mu}\right)^2 \right] dt, \tag{3.4}$$

and by introducing the new variable $y = \sqrt{2\mu}\left(t - \frac{f-f_0}{\mu}\right)$, we get:

$$\Psi(f) = \frac{1}{\sqrt{2\mu}} \exp\left[-j\frac{\pi}{\mu}(f-f_0)^2\right] \left[\int_{-Y_1}^{Y_2} \cos\left(\frac{\pi}{2}y^2\right) dy + j\int_{-Y_1}^{Y_2} \sin\left(\frac{\pi}{2}y^2\right) dy\right] \tag{3.5}$$

The new limits of the integral are:

$$Y_1 = \sqrt{2\mu}\left(\frac{T}{2} + \frac{f-f_0}{\mu}\right) \quad , \quad Y_2 = \sqrt{2\mu}\left(\frac{T}{2} - \frac{f-f_0}{\mu}\right) \ . \tag{3.6}$$

The limits can be written in an alternative form as a function of the time-bandwidth product D, where $D = TB$:

$$Y_1 = \sqrt{\frac{D}{2}}\left(1 + 2\frac{f-f_0}{B}\right) \quad , \quad Y_2 = \sqrt{\frac{D}{2}}\left(1 - 2\frac{f-f_0}{B}\right) \ . \tag{3.7}$$

The two integrals in (3.5) can be written as sum of the Fresnel integrals given by:

$$C(z) = \int_0^z \cos(\tfrac{\pi}{2}y^2)dy \quad , \quad S(z) = \int_0^z \sin(\tfrac{\pi}{2}y^2)dy \tag{3.8}$$

The final form of the frequency spectrum of the linear FM signal is then:

$$\Psi(f) = \frac{1}{\sqrt{2\mu}} \exp\left[-j\frac{\pi}{\mu}(f-f_0)^2\right] \{C(Y_1) + C(Y_2) + j\,[S(Y_1) + S(Y_2)]\} \tag{3.9}$$

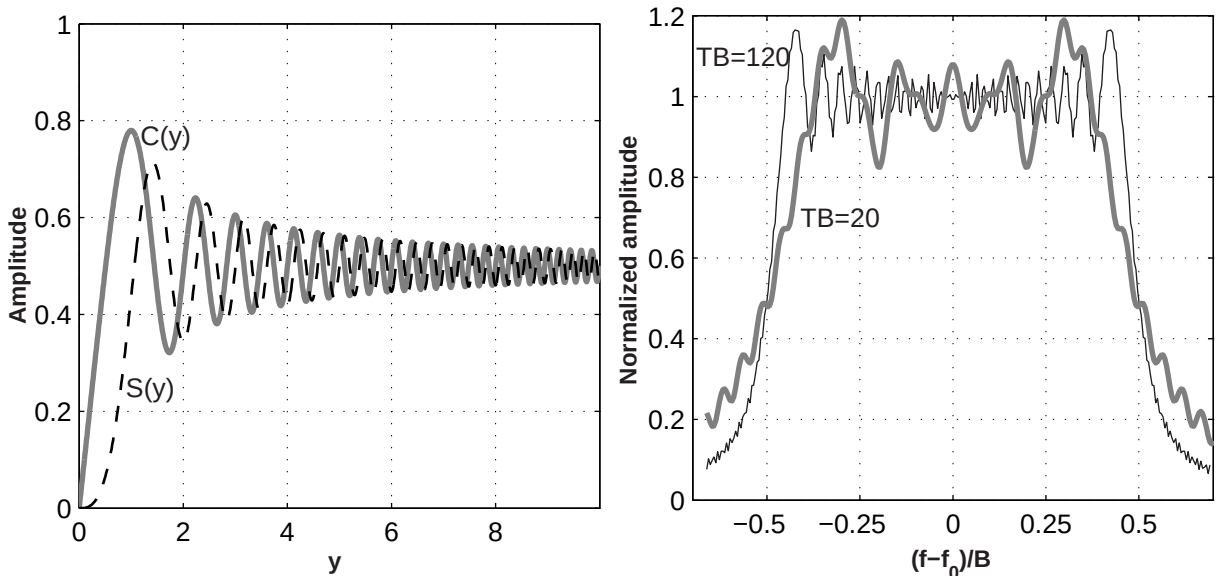

Figure 3.1: The Fresnel integrals (left) and the spectrum amplitude of the linear FM signal (right).

The Fresnel integrals defined in (3.8) are shown in Fig.3.1a. They approach the value $1/2$, when the argument is much larger than 1. The spectrum amplitude of the linear FM signal is plotted

in Fig.3.1b. Since the arguments of the Fresnel integrals are a function of the time-bandwidth product, the spectrum amplitude approaches the rectangular for high values of D. In particular, the amplitude of the oscillation decreases with D and the number of the ripples increases with D.

The phase $\vartheta(f)$ of the frequency spectrum from (3.9) can be written as a sum of two terms $\vartheta_1(f)$ and $\vartheta_2(f)$, where

$$\vartheta_1(f) = -\frac{\pi}{\mu}(f-f_0)^2 = -\pi D\left(\frac{f-f_0}{B}\right)^2 \quad , \quad \vartheta_2(f) = tan^{-1}\left[\frac{S(Y_1)+S(Y_2)}{C(Y_1)+C(Y_2)}\right] . \tag{3.10}$$

The term $\vartheta_2(f)$ is shown in Fig. 3.2. It is approximately constant and equal to $\pi/4$ within the signal bandwidth. The higher the time-bandwidth product, the better the approximation to a constant value. Thus, for high values of D, the phase of the signal is only the quadratic function of frequency given by $\vartheta_1(f)$ plus the constant $\pi/4$.

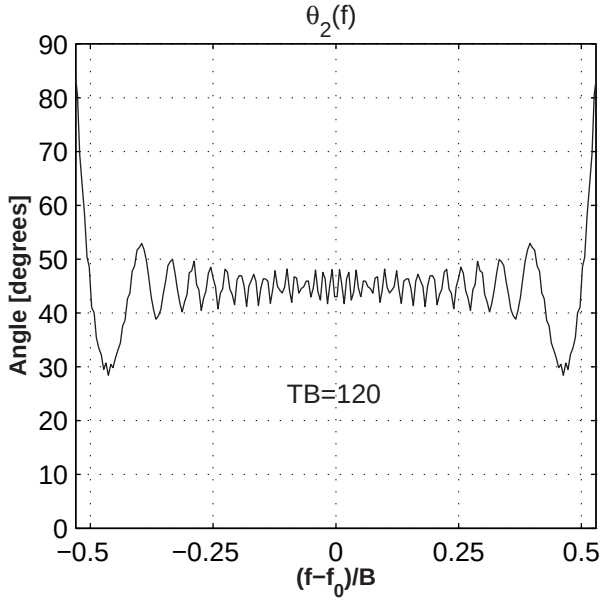

Figure 3.2: The phase distortion term $\vartheta_2(f)$ of the spectrum of a linear FM signal with a time-bandwidth product of 120.

3.3 Symmetry properties and their implications

The amplitude term $\left\{[C(Y_1)+C(Y_2)]^2+[S(Y_1)+S(Y_2)]^2\right\}^{1/2}$ is approximately equal to $\sqrt{2}$ for $f = f_0$. The Fresnel ripples oscillate around this value and for high TB, the amplitude can be considered rectangular in the passband with a value of $\sqrt{2}$. Rihaczek considers that the rectangular

approximation is valid when $D \geq 10$ [1]. The approximate phase of the spectrum is $\vartheta_1(f)+\pi/4$. Therefore, the high TB approximation of (3.9) for the spectrum of the linear FM signal is:

$$\Psi(f) = \frac{1}{\sqrt{\mu}}\text{rect}\left(\frac{f-f_0}{B}\right)\exp\left[-j\frac{\pi}{\mu}(f-f_0)^2+\frac{\pi}{4}\right] \tag{3.11}$$

Let the *group delay* $\tau_g(f)$ be defined as the derivative of the phase function $\vartheta(f)$ in the frequency domain. The group delay of the linear FM will then approximately be:

$$\tau_g(f) = \frac{1}{2\pi}\frac{d\vartheta(f)}{df} \simeq \frac{1}{2\pi}\frac{d\vartheta_1(f)}{df} = \frac{f_0-f}{\mu} \tag{3.12}$$

Using the concept of the group delay, the linear FM signal can be thought as one consisting of a number of adjacent spectral bands, each centered at a frequency f and delayed by the group delay $\tau_g(f)$. For $t=\tau_g$, (3.2) gives $f_i = f$, that is, the frequency f of each spectral band is the instantaneous frequency of the signal at a time equal to the group delay. In fact, it can be proven [1] that instantaneous frequency and group delay are inverse functions of each other. Using this relationship, Rihaczek [1] derived the approximate spectrum of the general linear FM signal with an arbitrary real envelope $a(t)$ given in (3.2), as:

$$\Psi(f) = \frac{1}{\sqrt{\mu}}a\left(\frac{f-f_0}{\mu}\right)\exp\left[-j\frac{\pi}{\mu}(f-f_0)^2+\frac{\pi}{4}\right]. \tag{3.13}$$

Therefore, the spectrum amplitude has the same function as the envelope of the signal in the time domain, simply being obtained by substituting t with $(f-f_0)/\mu$ and scaling the amplitude. Additionally, the phase functions $\phi(t)$ and $\vartheta(f)$ in the time and frequency domains respectively are both quadratic functions, or alternatively the instantaneous frequency and group delay are linear functions. There is, therefore, a symmetry in the time and frequency functions, which is a unique property of the linear FM signal.

The concepts of instantaneous frequency and group delay relate a given frequency with a corresponding delay, which can give a quantitative insight into the mechanism of pulse compression of the linear FM signal. The matched filter impulse response is the time reversed FM signal, as it was explained in Ch. 2. The group delay of the matched filter τ_g^F will be given by (3.12) with a minus sign. At one instance of time $t=t_a$, the signal has an instantaneous frequency $f_i(t_a) = f_0+\mu t_a$. The matched filter will delay this frequency by the group delay $\tau_g^F(f)$ for $f=f_i$. The time instant that this frequency appears at the matched filter output will then be:

$$t_a - \tau_g^F(f)|_{f=f_i(t_a)} = t_a - \frac{f_i(t_a)-f_0}{\mu} = t_a - \frac{(f_0+\mu t_a)-f_0}{\mu} = 0 \tag{3.14}$$

Therefore, all spectral components of the linear FM signal, independent of their frequency reach the filter output at the same time $t=0$. All frequencies are in phase at time zero and they add constructively, causing the output buildup.

When the returned signal is frequency shifted, the output will still be compressed because of the linearity of the group delay function. A frequency shift results in a shift in the group delay function without distorting its linear nature. The only consequence is merely a translation in time of the compressed output peak. This explains why compression of the linear FM signal is frequency shift invariant, a desired property for ultrasound imaging of attenuating media.

Another important consequence of the symmetry between time and frequency domains is the equivalence between time and frequency weighting. Weighting is a common technique to reduce the sidelobes of the compression output, and in the case of the linear FM signal, there is practically no difference if this weighting is applied on the time envelope of the filter or on its frequency spectrum.

However it should be remembered that the symmetry in time and frequency domains is only a high time-bandwidth product approximation. A rectangular envelope in the time domain actually yields a rectangular spectrum amplitude with Fresnel ripples as well as phase distortion. For high quality imaging, the Fresnel distortions have to be taken into account in the design of the transmitted signal and the appropriate matched filter. This will be discussed in the following chapter.

3.4 The matched filter response and the ridge ambiguity

The derivations of the analytic expression for the matched filter response and the ambiguity function of the linear FM signal are straightforward and can be found in almost all books from the radar literature. Therefore, the derivations will not be repeated here, and only the final expressions will be presented and discussed.

The complex envelope of the linear FM signal described by (3.1), when the real envelope is rectangular, is:

$$\mu(t) = \text{rect}\left(\frac{t}{T}\right) e^{j\pi(B/T)t^2} \tag{3.15}$$

The complex envelope can be substituted into (2.28) which defines the ambiguity function and is repeated here for convenience:

$$\chi(\tau, f_d) = \int_{-\infty}^{\infty} \mu(t) \cdot \mu^*(t-\tau) \cdot e^{j2\pi f_d t} dt. \tag{3.16}$$

The result for the real envelope of the ambiguity function is [34]:

$$|\chi(\tau, f_d)| = T \left| \frac{\sin\left[\pi D\left(\frac{\tau}{T}+\frac{f_d}{B}\right)\left(1-\frac{|\tau|}{T}\right)\right]}{\pi D\left(\frac{\tau}{T}+\frac{f_d}{B}\right)} \right| \tag{3.17}$$

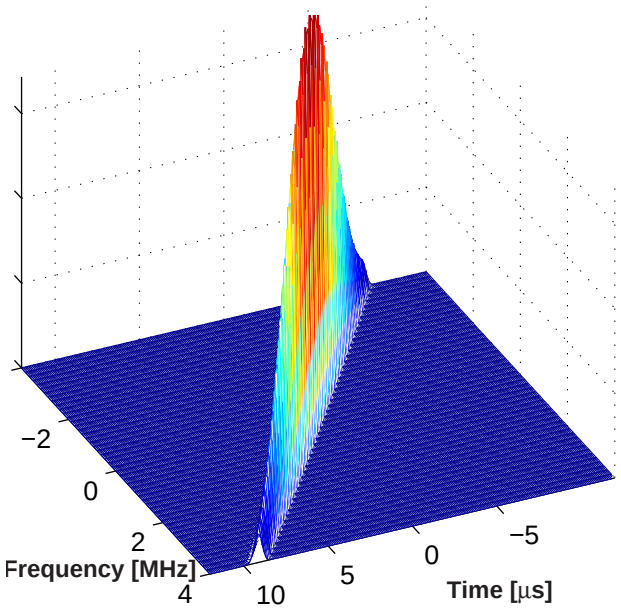

Figure 3.3: The ambiguity function of a linear FM signal with a time-bandwidth product of 140.

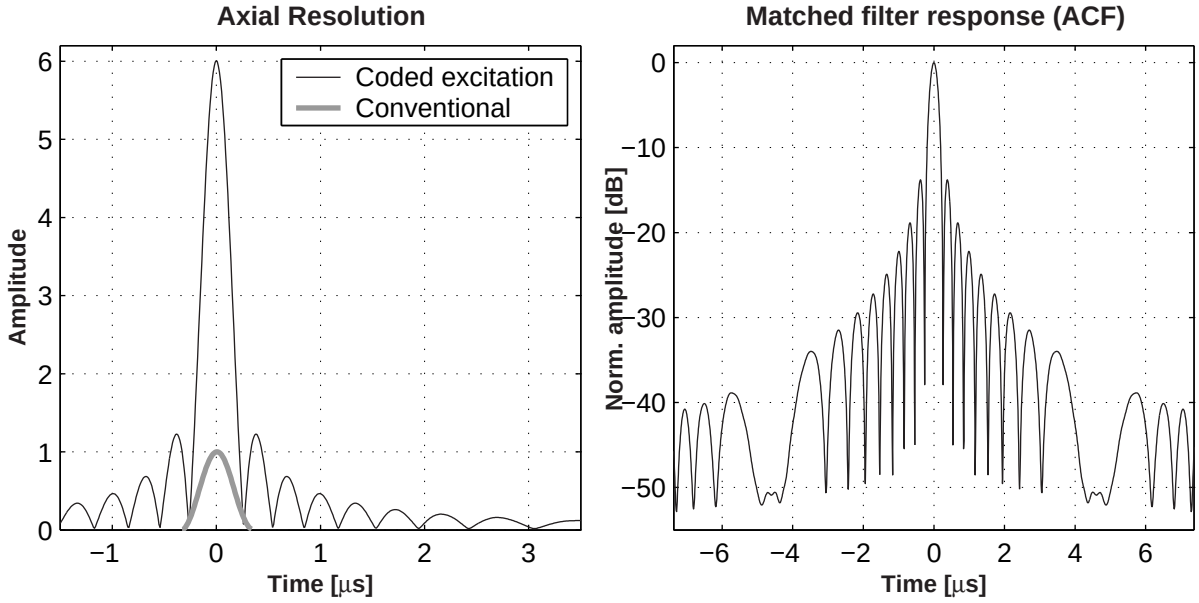

Figure 3.4: Resolution for pulsed and linear FM excitation (left graph). The pulse shown here (gray line) is the envelope of an apodized sinusoid of the carrrier frequency with Hanning apodization. The length is 2.7 cycles and is chosen to match the bandwidth of the chirp for direct comparisons. The black thin line is the compressed envelope of a linear FM signal with $D = TB = 36$. The same in logarithmic scale is shown in the right graph.

The ambiguity function for the linear FM signal is plotted in Fig. 3.3, while a contour plot was shown in the right plot of Fig. 2.5. The ambiguity surface is concentrated along the line $f_d = -(B/T)\tau$. The advantage of the ridge ambiguity function is that compression still occurs for a frequency mismatch. The amplitude drop of the ridge with increased frequency shifts is due to the incomplete overlap of signal spectrum and receiver passband, and can be avoided by simply widening the receiver passband [34]. The sidelobe behaviour along the τ and the f_d axes is very different. The sidelobe symmetry improves as TB increases.

Substituting (3.15) into the definition of the auto-correlation function (ACF) given in (1.17) we obtain [25]:

$$R_{\psi\psi}(\tau) = e^{j2\pi f_0 \tau} \int_{-\infty}^{\infty} \mu(t)\mu^*(t-\tau)dt = T \frac{\sin\left[\pi D \frac{\tau}{T}\left(1-\frac{|\tau|}{T}\right)\right]}{\pi D \frac{\tau}{T}} e^{j2\pi f_0 \tau} \tag{3.18}$$

Comparing (3.18) with (3.17), it can be seen what was pointed out in Chapter 2, i.e. that the matched filter output for zeros frequency shifts is the auto-correlation of the signal. Frequency modulation has disappeared, and the output is an amplitude-modulated signal at the carrier frequency f_0, with an approximately *sinc* shape. It is plotted in linear and logarithmic scale in Fig. 3.4. The approximation to the *sinc* function improves as the time-bandwidth product D increases.

The first zero is often taken as a measure of the time resolution. From (3.18), the first zero can be found [35]:

$$r_t = \frac{T}{2}\left(1 - \sqrt{1 - \frac{4}{BT}}\right) \approx \frac{1}{B} \tag{3.19}$$

Imaging with a short pulse of length T will yield axial resolution of $T = 1/B$. Thus, **axial resolution for a short (conventional) excitation pulse and for an FM-modulated (coded) excitation will be roughly the same, when the signals use the same bandwidth**. This is shown in Fig. 3.4.

The *compression ratio*, i.e. the ratio of the compressed pulse and uncompressed pulse widths is $T/(1/B) = TB$. The side effect of pulse compression on the linear FM is the resulting *sinc* sidelobes, not present in conventional pulse excitation.

3.5 Mismatched filtering

Range sidelobes represent an inherent part of the pulse compression mechanism. In an imaging situation, the effects of the time (or range) sidelobes extending on either side of the compressed pulse will be self-noise along the axial direction and masking of weaker echoes. For the linear FM signal, the highest of these sidelobes are the first ones, only -13.2 dB below the peak of the compressed pulse (Fig. 3.4b). The near sidelobes fall off at approximately 4 dB per sidelobe

interval, while the sidelobe null points are spaced approximately $1/B$ apart. The usual approach for sidelobe reduction is to apply a window function on the matched filter. Due to the symmetry properties weighting can be performed either in the time or in the frequency domain. Among a plethora of window functions, the Dolph-Chebyshev window exhibits the minimum main-lobe width for a specified sidelobe level and has been used extensively in radars [36]. We have used other windows such as the Blackman-Harris [37] windows, which yield narrow mainlobe width with adequately low sidelobe levels. Weighting techniques for the FM family of signals will be discussed in detail in the following chapter.

3.6 Gain in signal to noise ratio

The maximum of the ambiguity function for the linear FM signal (given in (3.17)) occurs at the origin and is equal to $T = 2E$. For a better appreciation of the peak value, the matched filter can be normalized in order to have unity gain at f_0. Then, the filter impulse response has a constant amplitude of $\sqrt{\mu} = \sqrt{D}/T$, and in this case, the peak value of the ambiguity function is $\sqrt{D}$. This represents the *gain factor* resulting from the pulse compression of a linear FM signal. In Fig. 3.4 a unity gain matched filter has been used, and the compression output has a peak value of $\sqrt{36}$=6 V. In this case, the expected gain in SNR will be $20\log(\sqrt{36}) = 15.5$ dB.

SNR degradation with mismatching

SNR is not drastically affected with weighting. The reason for that is that weighting acts on output values already attenuated at least 13.2 dB. Cook and Bernfeld [25] have calculated the mismatch loss relative to the matched filter for various weighting functions. To mention some of their results, a Hamming window which reduces the sidelobes to -42.8 dB has a mismatch loss of 1.34 dB, while a Dolph-Chebyshev window which reduces the sidelobes to -40 dB has a mismatch loss of 1.2 dB.

3.7 Non-linear FM modulation

The non-linear FM waveform has a modulation function which is designed to provide a desired amplitude spectrum. Such design can allow, for instance, to generate a signal that matches the spectrum of an ultrasonic system, while at the same time maintains the rectangular (high-energy) time-domain envelope. Non-linear FM signals have a spectrum weighting function inherently in their modulation function, which offers the advantage that a pure matched filter gives low side-lobes. Thus, the loss in signal-to-noise ratio associated with weighting by the usual mismatching technique is eliminated.

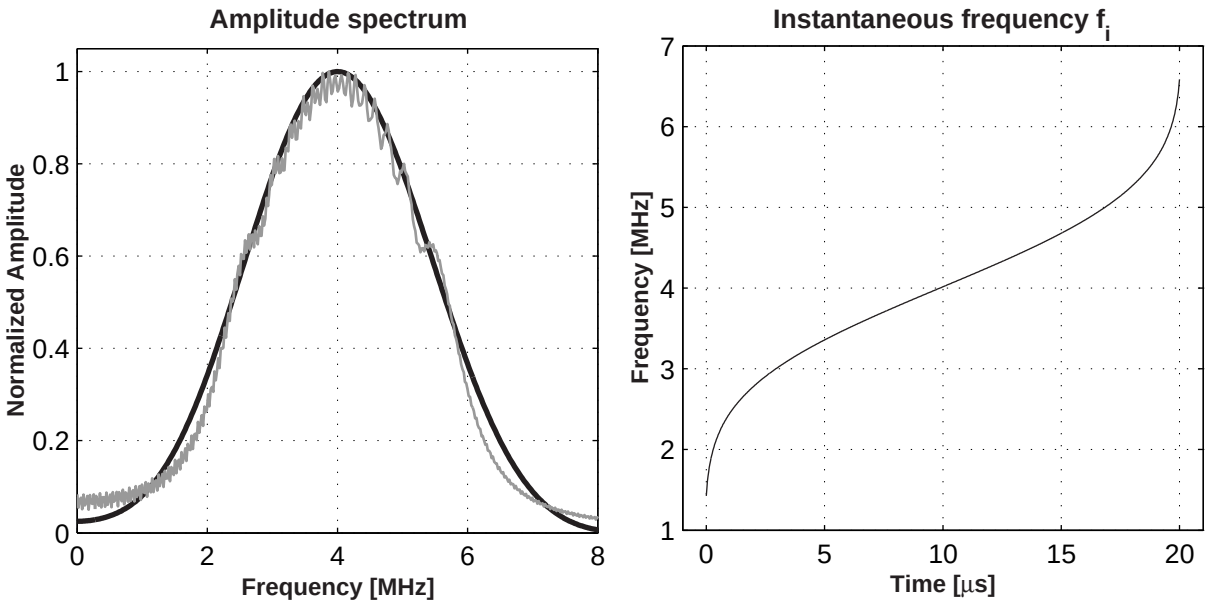

Figure 3.5: Non-linear FM signal design. The signal is designed to match the amplitude spectrum of a simulated transfer function of an ultrasonic transducer (left graph, black line). The amplitude spectrum of the resulting signal (gray line) is very close to the one specified. The instantaneous frequency of the signal is the non-linear sigma-shaped function shown at the right.

The fact that the instantaneous frequency of the linear FM is a linear function of time means that all spectral components are swept for the same time duration, which results in a spectrum with constant amplitude. Intuitively, a bell-shaped amplitude spectrum will correspond to an instantaneous frequency function which remains for longer time in the center of the spectrum and less time at the edges. The mathematical design of such instantaneous frequency function is based on the following equation, which actually relates the group delay function with the amplitude of the spectrum [18]:

$$-\frac{d\tau_g(f)}{df} = K \cdot |\Psi(f)|^2 \quad \Rightarrow \quad \tau_g(f) = \int_0^f C_1 \cdot |\Psi(f)|^2 + C_2 \tag{3.20}$$

where the integration constants C_1 and C_2 are chosen so that $\tau_g(f_1) = 0$ and $\tau_g(f_2) = T$. These equations are only approximations and do not include the actual Fresnel ripples of the spectrum.

The steps for designing a non-linear FM signal are the following:

- Design a weighted amplitude spectrum. The transfer function of the ultrasonic transducer is the best choice in terms of signal-to-noise ratio of the ultrasound echoes [18], but other windowing functions such as the Dolph-Chebyshev window might yield signals with lower sidelobes.

- Perform numerical integration of the amplitude spectrum according to (3.20). The resulting group delay function increases monotonously between the frequencies $f_1 = f_0 - \frac{B}{2}$ and $f_2 = f_0 + \frac{B}{2}$.

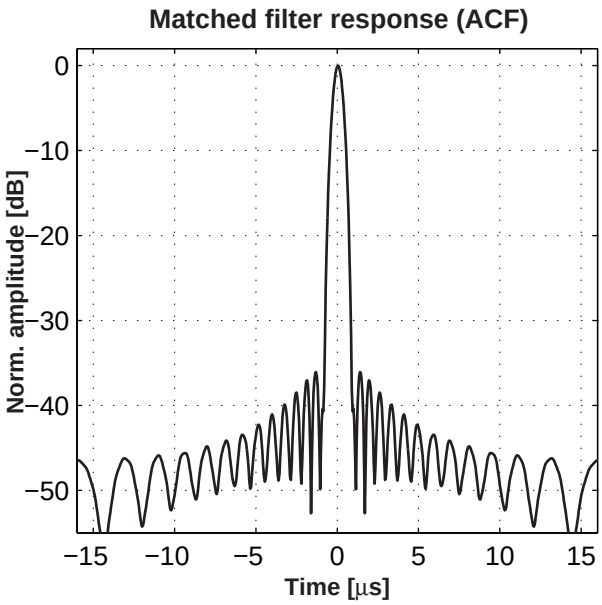

Figure 3.6: The auto-correlation function of the non-linear FM signal. The sidelobes are low without any weighting.

- Calculate the inverse of the group delay function, which will be the instantaneous frequency function of time. This is done by exchanging the abscissa and the ordinate. Since the abscissa values of the group delay are irregularly spaced, intermediate values are calculated using linear or spline interpolation.

- Perform numerical integration of the instantaneous frequency and construct the FM signal as:

$$x(t) = \sin\left(2\pi \int_0^t f_i(\tau) d\tau\right) \quad , \quad 0 \leq t \leq T \tag{3.21}$$

The results of such a design of a non-linear FM signal of length 20 μs are shown in Fig. 3.5. The amplitude spectrum of the resulting signal (gray line) is very close to the one specified, apart from the Fresnel ripples. Note that the methods for reduction of the Fresnel ripples that will be discussed in the following chapter for linear FM signals are applicable here. The auto-correlation function (matched filter response) shown in Fig. 3.6 shows that this design yields sidelobes down to -36 dB without any weighting. The gain in SNR for this FM signal is 17.5 dB, while the -20 dB axial resolution (mainlobe width) is 1.17λ.

CHAPTER

FOUR

Weighting of FM signals and sidelobe reduction for ultrasound imaging

4.1 Weighting in time and frequency domain

As it was shown in the previous chapter, the matched filter output of the linear FM signal has an approximate *sinc*-like envelope with sidelobes only -13.2 dB below the peak independently on the time-bandwidth product. It is well established from the radar literature, that these sidelobes can be reduced by weighting either in the frequency or in the time domain. If the spectrum of the linear FM was strictly rectangular and its phase was strictly quadratic, i.e. if there were no Fresnel distortions in either the spectrum amplitude or the phase, the two weighting methods would be identical. In the presence of ripples, time and frequency weighting are only approximately equivalent, with time weighting having slightly better performance than frequency weighting for time-bandwidth products less than about 200 [38].

The response for a matched filter given by (2.6) and (2.11) is repeated here for convenience:

$$\gamma(t) = \int_0^\infty \Psi(f) \cdot H(f) e^{j2\pi ft} df = \int_0^\infty |\Psi(f)|^2 e^{j2\pi ft} df, \quad where \quad H(f) = \Psi^*(f) \ . \tag{4.1}$$

Frequency-domain filtering is the shaping of the amplitude spectrum of the signal with a real window function $W(f)$. The output of the so called *mismatched filter* is then:

$$\gamma_w(t) = \int_0^\infty \Psi(f) \cdot H_w(f) e^{j2\pi ft} df = \int_0^\infty W(f) \cdot |\Psi(f)|^2 e^{j2\pi ft} df, \quad where \quad H_w(f) = W(f) \cdot \Psi^*(f) \ . \tag{4.2}$$

Using the rectangular FM spectrum approximation, the function $W(f)$ can be designed as the Fourier transform of a desired output time function $\gamma_w(t)$:

$$H_w(f) \ = \int \gamma_w(t) e^{-j2\pi ft} dt \ . \tag{4.3}$$

With this technique, the mismatched condition applies only on the amplitude spectrum, while the receiver transfer phase remains conjugate function of the spectrum phase. The *sinc* function of the matched filter response corresponds to the rectangular window.

Time-domain filtering is based on amplitude weighting of the envelope of the transmitted FM signal or of the matched filter impulse response. The former case is usually avoided, because it is desirable to transmit as much energy as possible. In the latter case, the transfer function of a mismatched filter using time-weighting is the complex conjugate of an amplitude-weighted version of the excitation signal. Time weighting filter design is based on the following relationships:

$$\begin{aligned} FM \quad signal: &\quad \psi(t) \overset{F}{\longleftrightarrow} \Psi(t) \\ Mismatched \quad filter: &\quad \psi(t)\cdot w(t) \overset{F}{\longleftrightarrow} H(t) \\ Compression \quad output: &\quad \gamma(t) \overset{F}{\longleftrightarrow} \Psi(t)\cdot H^*(t) \end{aligned} \tag{4.4}$$

4.2 Weighting functions and tapering

Sidelobe reduction in matched filter receivers is reduced to the choice of an appropriate window function. There is a plethora of window functions reported, and Harris [37] gives a thorough review of the most common windows and their properties. Windows such as Hanning, Kaiser, Blackman, Hamming etc. are discussed.

A more systematic window design optimization technique is reported in an excellent paper by Adams [31]. It is based on a trade-off between the window design parameters, which are the mainlobe width, the total sidelobe energy and the peak sidelobe level. The energy contained in the sidelobe region is associated with the contrast in the image. High sidelobe energy will cause leakage of energy from bright areas into dark areas in the image. Such case is imaging of cysts, where an hypoechoic dark region is surrounded by a bright contour. The window function that minimizes the sidelobe energy is the prolate-spheroidal window. However the resulting peak sidelobe level can be very high.

On the other extreme, stands the Dolph-Chebyshev window, which minimizes the peak sidelobe level. The Dolph-Chebyshev window exhibits the minimum mainlobe width for a specified constant sidelobe level and has been used extensively in radar systems. Adams describes a method for optimal windows that lie between these two extremes.

Fig. 4.1 gives two examples of mismatched time-weighting with a Hamming and a Dolph-Chebyshev window applied. The matched filter response is also plotted for comparison. A FM signal with a time-bandwidth product of 62 has been used. For the Hamming window the first sidelobes close the mainlobe have been reduced to about -40 dB. The deviation from the theoretical -42.8 dB as well as the lack of symmetry are an indication that the limited-energy relatively short FM signal is not band-limited and therefore not equal to the analytical signal. This can be as-

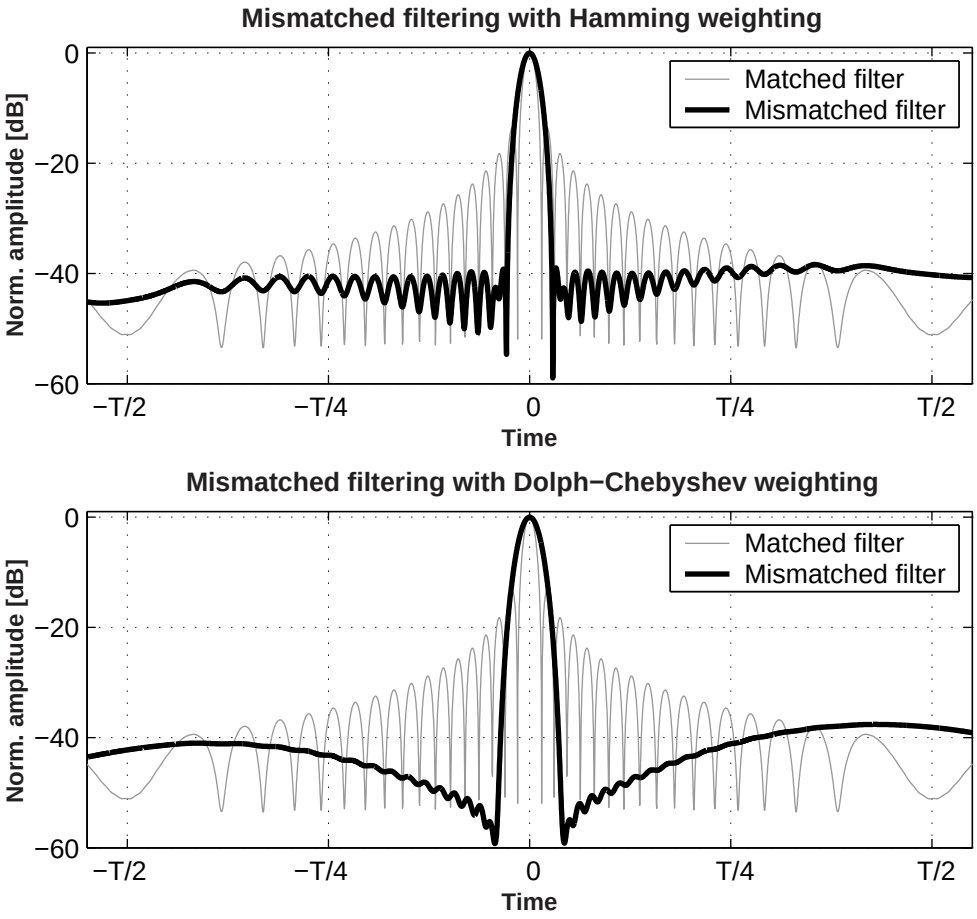

Figure 4.1: Compression outputs for two mismatched filters based on time-weighting with Hamming (upper graph) and Dolph-Chebyshev windowing (lower graph).

sessed by comparing the deviation of the spectrum of the signal with the spectrum of the analytical signal given by (3.9). Further reduction of the first sidelobes is achieved with the Dolph-Chebyshev window designed to yield -60 dB constant sidelobe level. The lower sidelobe levels achieved with mismatched filtering are accompanied with a) a broadening of the mainlobe and associated loss in axial resolution, and b) a loss in SNR. The -20 dB mainlobe width (axial resolution) of the mismatched filter output using Hamming weighting is 1.8λ, and the mainlobe width for Dolph-Chebyshev weighting is 2.2λ. For comparison, the mainlobe width for the matched filter case is 1.1λ and for a conventional 4-cycle pulse with Hanning apodization is 1.2λ.

4.3 The effect of the ultrasonic transducer on pulse compression

In ultrasound, bandpass filtering of the excitation signal occurs inherently from the transducer. This has two consequences: a) the gain in SNR is significantly lower than what is achieved in radar systems, i.e. lower than $\sqrt{TB}$, and b) the sidelobes from pulse compression is less of a problem.

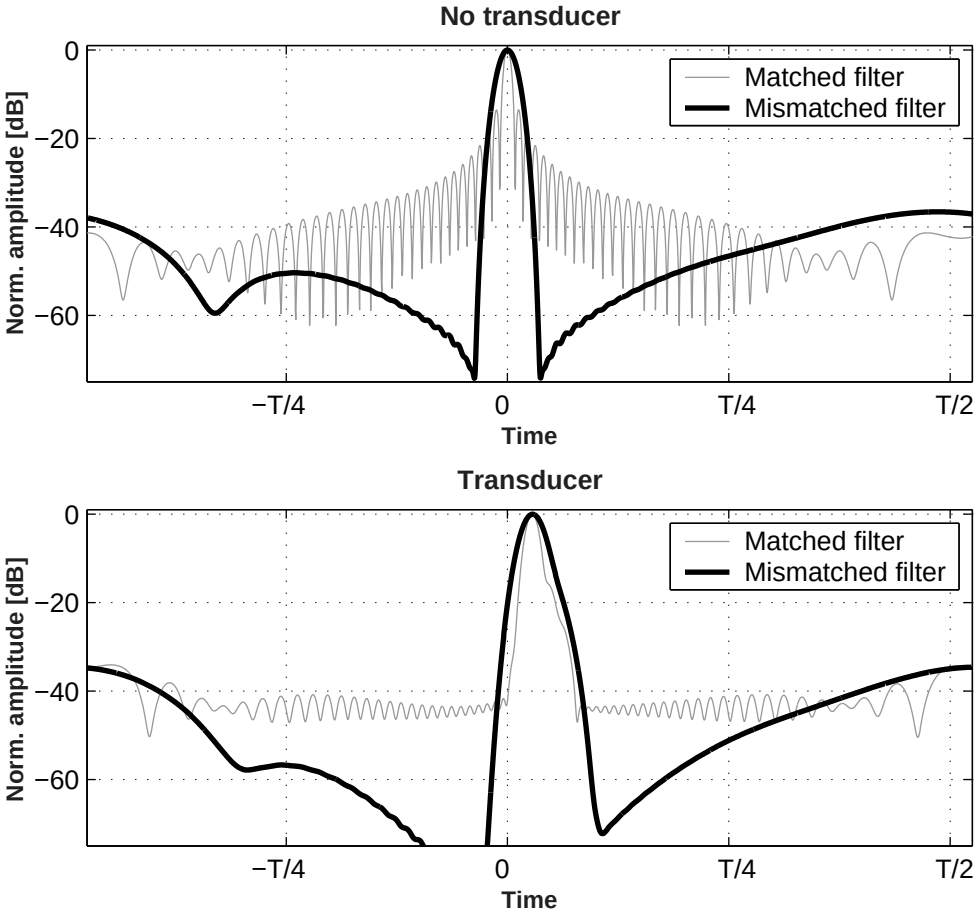

Figure 4.2: The effect of the transducer on pulse compression. The faint lines are the matched-filter outputs and the bold lines are the outputs when a Dolph-Chebyshev window has been applied to the compression filter. The specified sidelobe level for the window was -90 dB.

The transducer effect on pulse compression is illustrated in Fig. 4.2. The FM signal is convolved with the measured pulse-echo impulse response of an ultrasonic transducer and the resulting signal is used as the input in the matched filter. The transducer used is a 4 MHz single-element transducer (B-K Medical) and the calculated -6 dB fractional bandwidth was 65%.

When the bandwidth of the chirp matches the transducer's bandwidth, the presence of the transducer reduces the near range sidelobes from -13.2 to -32 dB below the mainlobe. The linear FM signal used for Fig. 4.2 had 1.2 higher bandwidth, and that reduces the sidelobes further down to -42 dB. When mismatched filtering is applied (bold lines of Fig. 4.2), near range sidelobes can be further reduced at the expense of an additional broadening in the main lobe. Further reduction of the near sidelobes is possible when the compression filter is matched not to the excitation signal but to the signal after that has passed through a simulated transducer. However the latter approach shapes the filter according to the transducer's bandwidth and is less stable to deviations of the simulated impulse response to the actual one. The effect of the transducer is equivalent to an additional weighting that, although suboptimal, is beneficial to sidelobe reduction. Fig. 4.2b shows that matched filtering of linear FM signals transmitted by an ultrasound transducer yields compression with sidelobe levels of -37 dB without any weighting on the receiver. Transmitting a weighted pulse (which occurs by the transducer) is desirable, because a weighted transmitted signal or spectrum is less sensitive to mismatches in frequency shifts [39]. In high power radars, weighting of the transmitted pulse is generally not used, since the final amplifier stages operate class C and are not subject to amplitude control [25].

4.4 Fresnel ripples and paired-echoes sidelobes

Fig. 4.1 shows that weighting reduces the sidelobes which are close to the mainlobe, but not the distant sidelobes which are in the ±T/2 region. The distant sidelobes increase the mainlobe-to-sidelobe energy ratio (MSR) and, thus, degrade the image quality. As it will be shown, these sidelobes are not controlled by the weighting function. The efficiency of windowing is based on the concept of shaping a rectangular spectrum in order to smooth out the sidelobes of the *sinc* function in the time domain, in accordance with (4.3). This equation is only an approximation, which does not take into account the Fresnel ripples of the spectrum amplitude and phase described by (3.9) and (3.10). The distant range sidelobes are paired-echo distortions due to deviation from the ideal rectangular spectrum amplitude and quadratic phase. The ripples given by the Fresnel integrals of (3.9) can be approximated with a combination of sinusoids [25]. With such an approximation, calculations with the nonclosed form of the Fresnel integrals are avoided, and useful results can be obtained. The effect of amplitude error in the compression output can be assessed by the following Fourier pairs [36]:

$$\begin{array}{ccc} G(f) & \overset{F}{\longleftrightarrow} & g(t) \\ G(f)\left[1+a_n\cos\left(2\pi n\frac{f}{B}\right)\right] & \overset{F}{\longleftrightarrow} & \frac{a_n}{2}g\left(t+\frac{n}{B}\right)+g(t)+\frac{a_n}{2}g\left(t-\frac{n}{B}\right) \end{array} \tag{4.5}$$

The presence of n spectrum amplitude ripples of amplitude a_n over the passband B of a signal spectrum $G(f)$ create symmetrical paired echoes in the time domain delayed and advanced from the main signal by n/B and scaled in amplitude by $a_n/2$. Similar paired-echoes result from the ripples in the spectrum phase [36]. Fig. 4.3 shows an application of (4.5) for a spectrum having 20

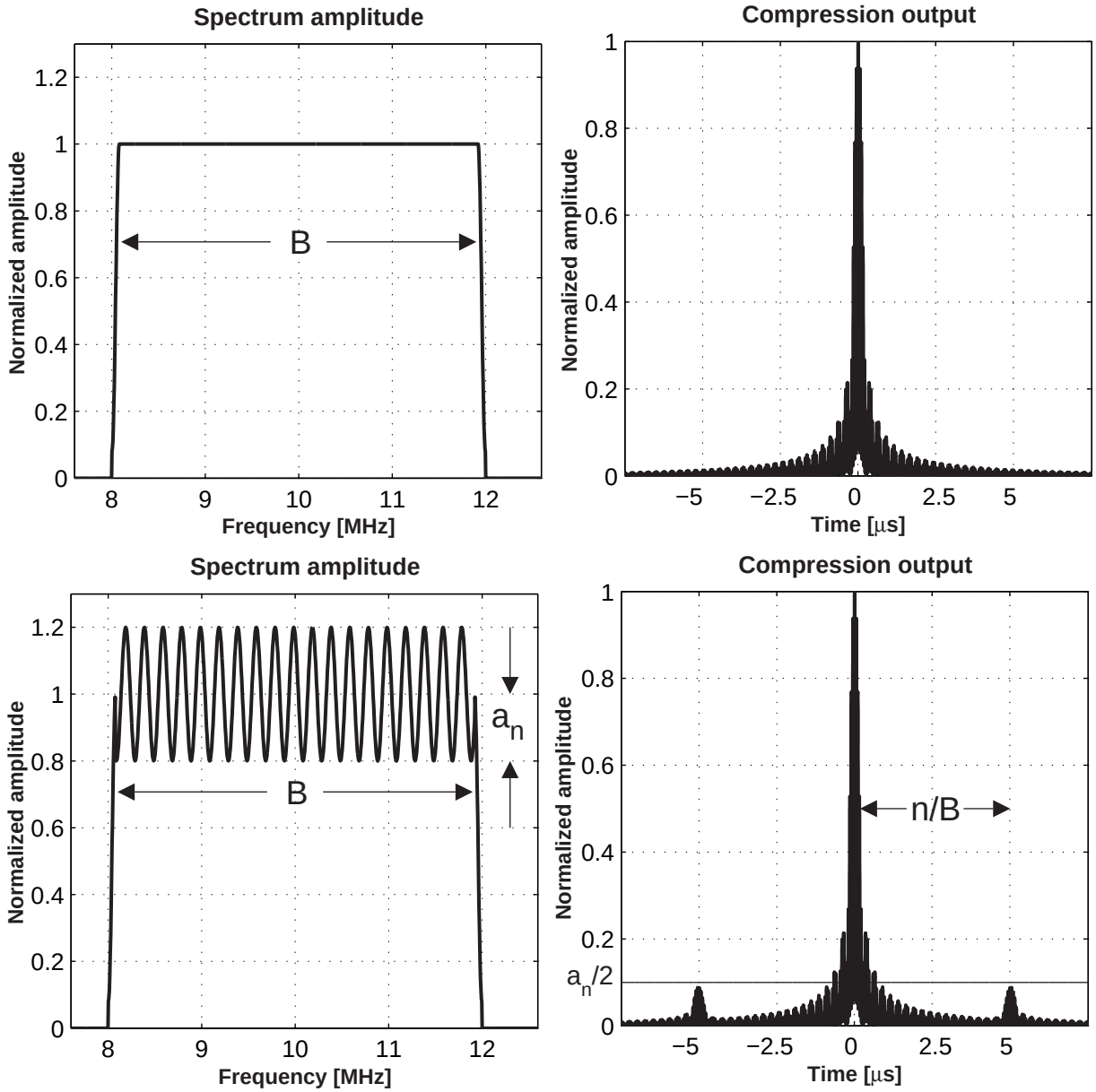

Figure 4.3: The presence of n spectrum amplitude ripples of amplitude a_n over the passband B of a signal spectrum $G(f)$ create symmetrical paired echoes in the time domain delayed and advanced from the main signal by n/B and scaled in amplitude by $a_n/2$.

amplitude ripples (n=20) which oscillate 20% about the constant value of the spectrum amplitude (a_n=0.2). In this case, the resulting distant sidelobes have a relative amplitude of $20\log(a_n/2)$=20 dB below the mainlobe peak. The number of ripples and their oscillating amplitude is a function of the time-bandwidth product. Therefore, the amplitude of the resulting sidelobes and their time displacement will also be a function of TB. For an FM signal with TB=100, the approximate paired-echo sidelobes are at about $\pm$T/2 relative to the center of the compressed pulse with an amplitude of about -35 to -40 dB [40]. The displayed dynamic range of a typical ultrasound image is 50 dB or greater, and therefore methods that compensate for the distant axial sidelobes have to be employed. This is an important consideration in the design of efficient coded excitation systems with low sidelobes.

4.5 Amplitude and phase predistortion

In order to eliminate the paired-echo distortions, frequency weighting functions could be constructed that would cancel the spectrum ripples of the linear FM signal. The ripple structure of the weighting function to be designed is a function of the time-bandwidth product and, thus, different weighting functions should be designed for various TB products. There are two drawbacks of such attempt: the transmitted FM signal passes through a transducer with an unknown impulse response, before it starts propagating through the tissues. This will cause distortions and the ripple cancellation of the weighting function might not be exact. The second drawback is that the sidelobe structure will become sensitive to frequency shifts of the returned signal.

A more practical and robust approach is to search for methods to reduce the ripples of the FM spectrum, and then apply a smooth weighting function on the receiver [25]. Three different approaches for ripple reduction have been tested by our group [20]: envelope modulation, phase predistortion, and time weighting of the edges.

Envelope modulation

This approach is based on the symmetry property of the linear FM signal in time and frequency. A linear FM signal with a constant amplitude envelope in the time domain yields a rectangular amplitude spectrum with ripple distortion. The reciprocal signal with a constant amplitude spectrum and strictly quadratic phase will have an amplitude and phase in the time domain given by Eq.(3.9), if T is replaced with B and t with f [1]. Using this concept, an FM signal with duration of $T = 20\mu s$ and a total bandwidth of 5.7 MHz ($TB = 114$) has been designed. The time signal and its spectrum amplitude are shown in Fig. 4.4. Although there is some improvement, there are still substantial ripples in both the amplitude and the phase of the spectrum (not shown here). The technique, however, is very efficient for small fractional bandwidths, when the exponential signal

is analytic and Eq.(3.9) is a good approximation for the signal spectrum. The disadvantages of this technique are a) the time signal is complex, b) such amplitude-modulation requires amplifiers with fast response, and c) the efficiency is low for time-bandwidth products greater than 30.

Phase predistortion

An alternative approach with similar performance is to use a phase predistortion of the transmitted signal, since amplitude and phase distortions have functional similarity and produce similar effects [40]. Both quadratic and cubic predistortion functions have been previously reported [40, 38]. In both papers, the quadratic phase function of the signal has been modified at the beginning and at the end, approximating essentially a non-linear FM instantaneous frequency function using linear [40] or quadratic [38] segments. The design parameters are the length of the two end segments and the slopes of their instantaneous frequency curves. Similar to non-linear FM signals, phase predistorted chirps have the advantage of sending out more energy, but are more vulnerable to phase distortions added from acoustic propagation and frequency shifts.

Amplitude tapering

The amplitude ripples of the spectrum can be attributed to the sharp rise-time of the time envelope, since a linear FM signal with infinite duration has no ripples. Data correlating pulse rise-time and spectrum ripple have been previously reported [40]. Based on this analysis, a modified chirp with amplitude tapering of the transmitted signal has been generated. The attainable ripple reduction is a function of the signal's bandwidth, the choice of the tapering function and its duration. Fig. 4.5 shows the effect of amplitude tapering on spectrum ripple reduction. The tapered function used in Fig. 4.5 is a Tukey window with a duration of $0.15 \cdot T$.

4.6 Proposed excitation/compression scheme

As it was shown in the previous section, amplitude tapering is the most efficient way to reduce the Fresnel ripples of the spectrum, if the power amplifier allows control of the transmitted pulse rise time. For a given duration T, the design parameters are:

i) the frequency band that is swept relative to the transducer's bandwidth

ii) the choice of the tapering function and its duration

iii) the choice of the weighting function.

The transducer's bandwidth can be modeled, but the phase of its transfer function is unknown and, thus, can not be compensated for. The effort here is to minimize the effect of the convolution between the transducer impulse response and the excitation signal on the design of the waveforms and the compression filter. This is done in the excitation signal by sweeping a bandwidth slightly larger than the transducer's frequency passband. This has an additional advantage: the effect of the two main rise and fall overshoots of the spectrum (see Fig. 4.5) is minimized. The increased bandwidth of the excitation waveform also yields a better axial resolution and higher gain. With this design, the range sidelobes can be as low as -88.2 dB (Fig. 4.6a). The second graph of Fig. 4.6 shows the compression output, when the filter is matched to the tapered signal convolved with a simulated transducer impulse response. This approach broadens the mainlobe but makes the sidelobe performance completely independent on the actual impulse response of the transducer. Fig. 4.7 shows optimized compression outputs for different frequencies and signal duration. Higher time-bandwidth product tapered FM signals have nearly ideal rectangular spectrum, allowing efficient sidelobe reduction well below 100 dB by the weighting function. If the length of the excitation signal is doubled, the sidelobes drop down to -105.5 dB with a good resolution of 1.45λ (Fig. 4.7). For comparisons, the -20 dB axial resolution with a typical pulse excitation is 1.46λ. The design gives even better results for higher frequencies.

The appropriate choice of the weighting function is a tradeoff between axial resolution and sidelobe levels. This is illustrated in Fig. 4.8, where the compression resolution and peak sidelobe level of 15 compression filters with different Dolph-Chebyshev window functions are shown. An implementation of the new scheme can easily give the flexibility to switch between finer resolution or lower range sidelobes, depending on the application needs. Dolph-Chebyshev windows, although optimal in sidelobe level performance, exhibit spikes at their leading and trailing edges. Utilization of other optimal windows described in [31] might result in better performance.

The compression for the design shown Fig. 4.6b retains its characteristics regardless of the actual transducer impulse response. For the more sensitive design of Fig. 4.6a, the effect of the transducer is shown in (Fig. 4.9). The degradation in the compression properties is still not significant for this design.

In conclusion, a new FM-coded scheme has been presented with distinct features, that make it attractive to the implementation of high-performance coded ultrasound systems. It is based on a predistorted FM signal and optimized compression filter. The range resolution, that can be achieved, is comparable to that of a conventional system. The range sidelobes, at the same time, are well below the limits of the typical dynamic range of an ultrasound image. The energy of the sidelobe region is also reduced by lowering the distant sidelobes caused by the ripples of the spectrum's amplitude. Evaluation of this coded excitation scheme through simulations and clinical images will be presented in Chapters 6 and 7 respectively.

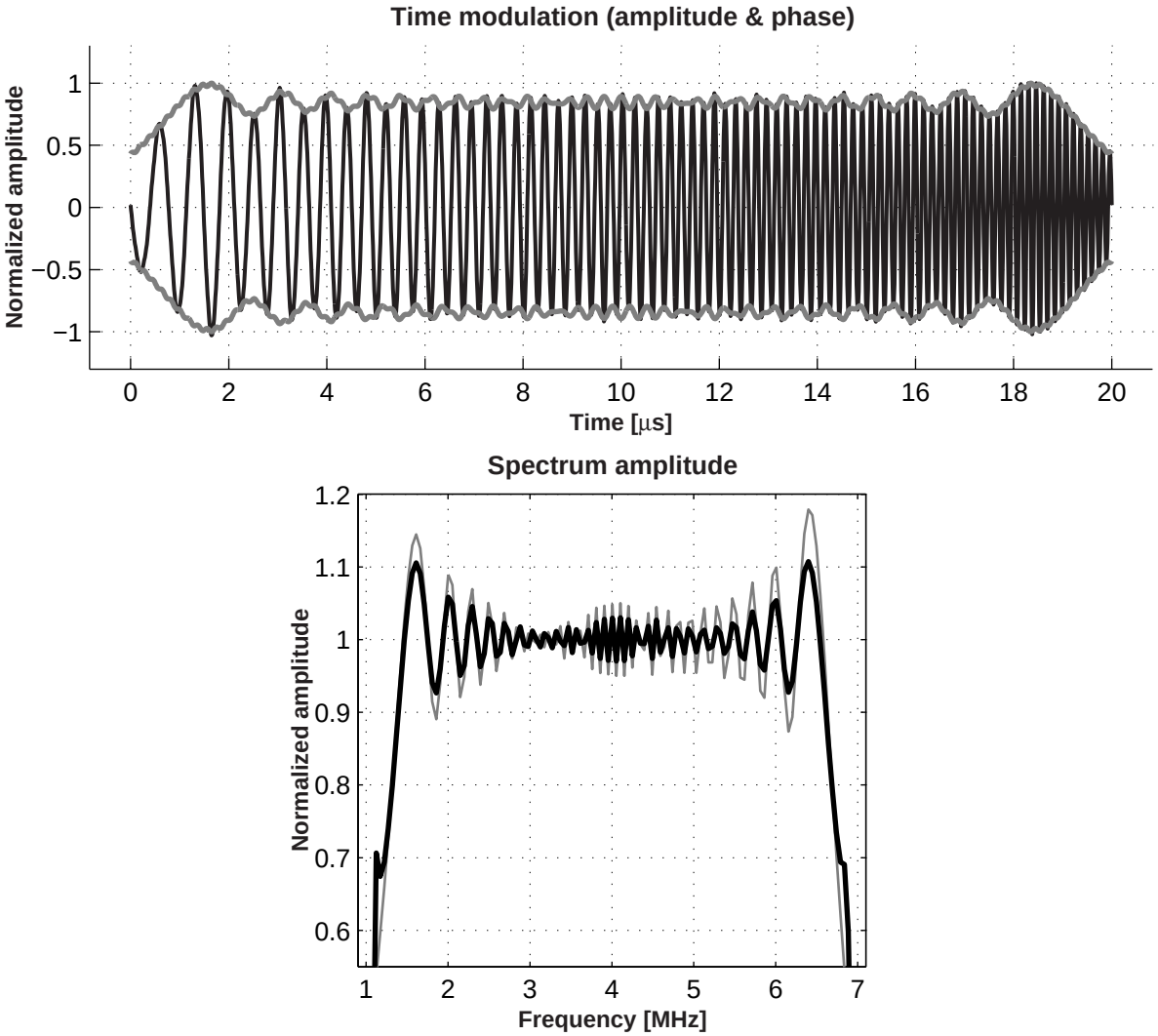

Figure 4.4: FM signal with Fresnel distortion in amplitude and phase (up), and its spectrum amplitude (down). The spectrum of a linear FM signal with constant amplitude envelope is shown for comparison in gray in the bottom graph.

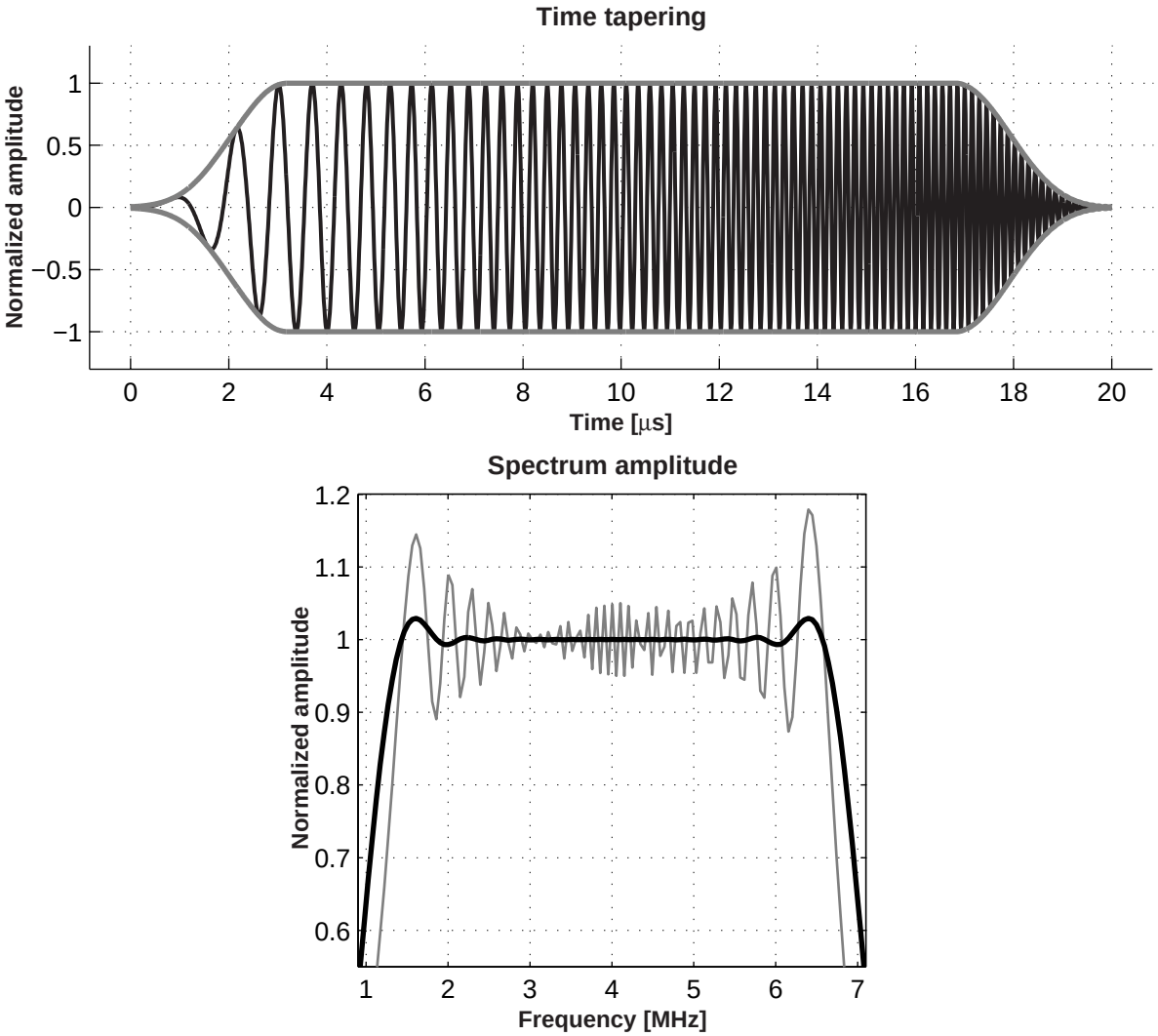

Figure 4.5: FM signal with amplitude tapering of the edges (up), and its spectrum amplitude (down), showing susbstantial ripple reduction. The spectrum of a linear FM signal with constant amplitude envelope is shown for comparison in gray in the bottom graph.

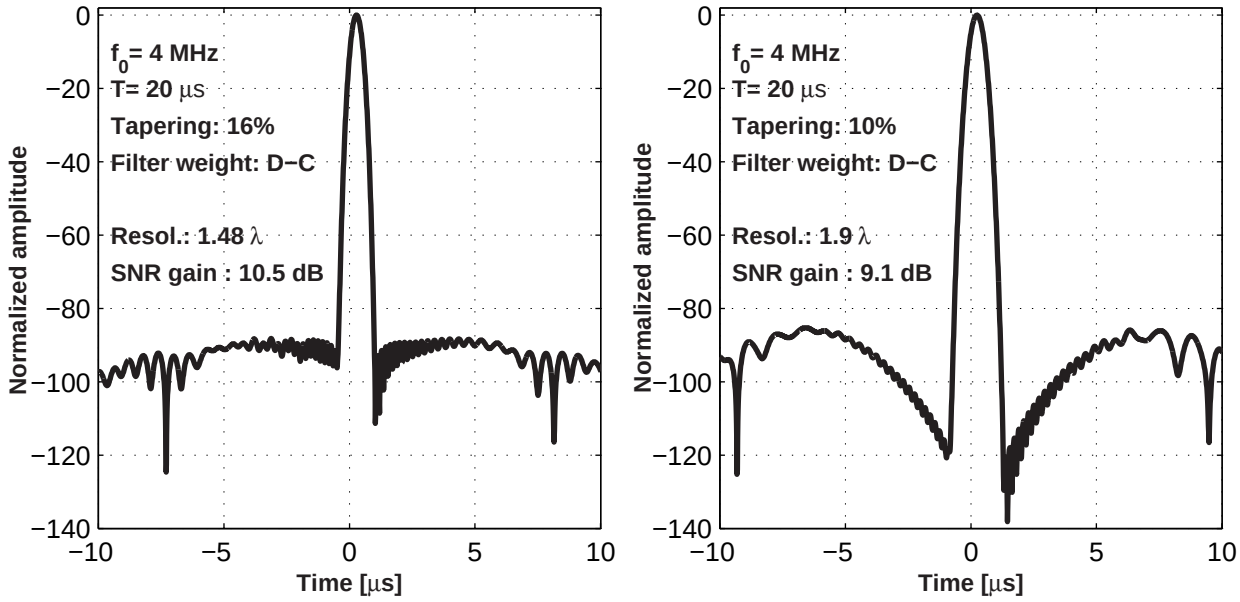

Figure 4.6: Optimized compression outputs for FM signals with amplitude tapering. The first scheme uses a weighted filter matched to the tapered signal, while in the second the filter is matched to the signal convolved with a simulated transducer impulse response.

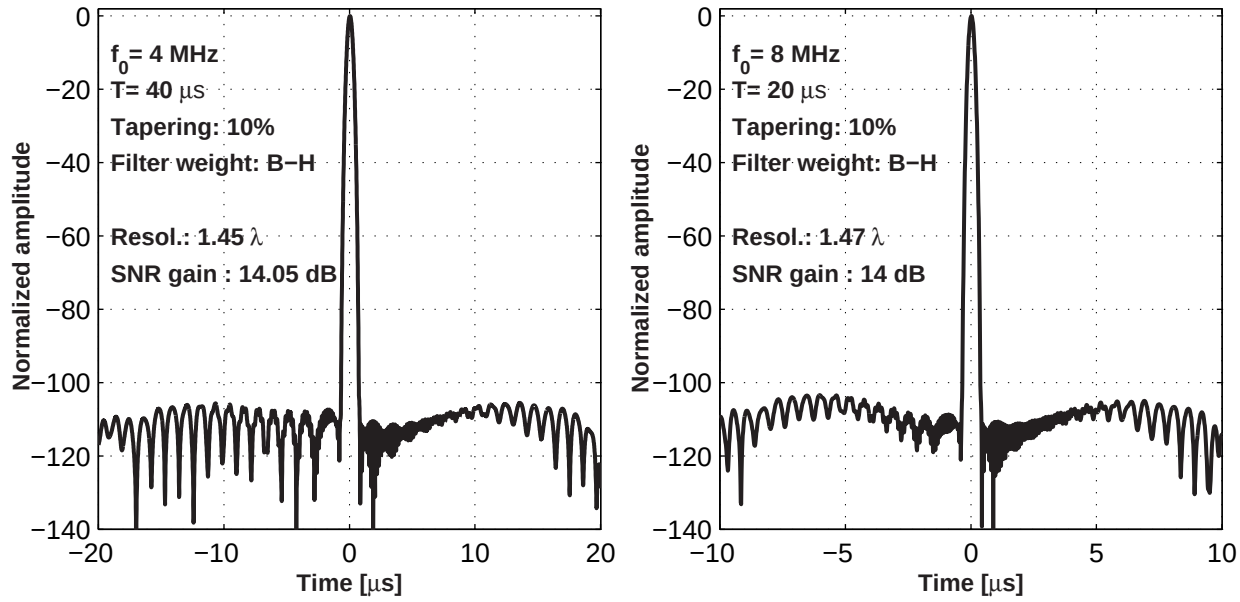

Figure 4.7: Optimized compression outputs for FM signals with amplitude tapering for different carrier frequencies and signal duration.

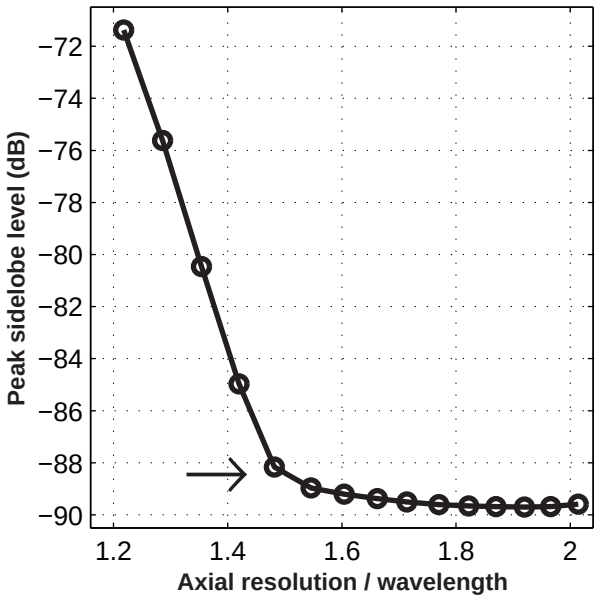

Figure 4.8: Trade-off between sidelobe level and axial resolution for a number of Dolph-Chebyshev mismatched compression filters. For most applications, the appropriate choice is at points in the knees of the curve as the one indicated by the arrow.

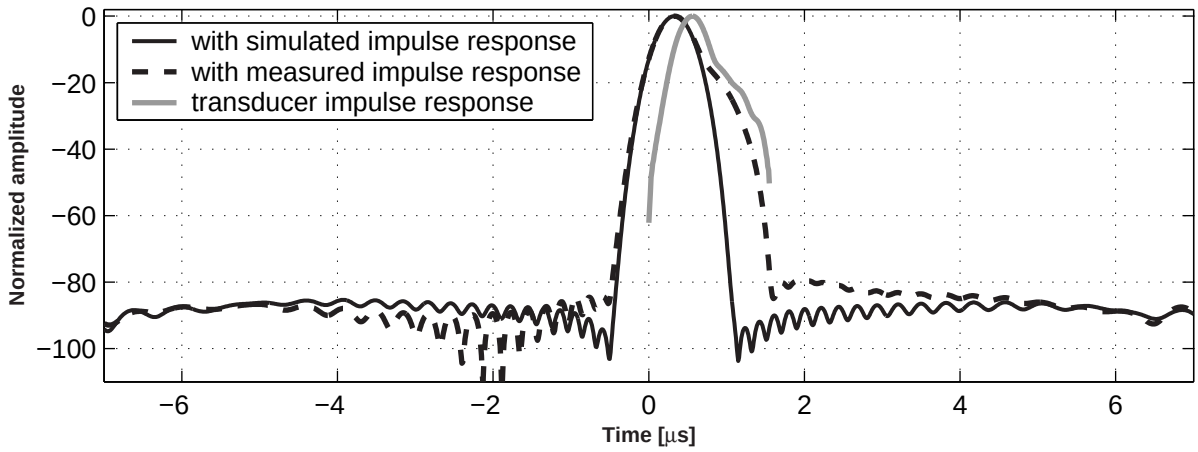

Figure 4.9: The effect of the transducer on the new scheme. The black line is the compression output of Fig. 4.6a. The dotted line is the compressed output when the actual transducer is used. It is shaped by the envelope of the measured impulse response shown by the gray line.

CHAPTER

FIVE

Phase-modulated signals

From the discussion in Chapter 2, it is clear that coded waveforms with ambiguity function sensitive to frequency shifts can only have limited use in ultrasound imaging. These waveforms include the majority of the binary codes (Barker codes, pseudo-random and m-sequences). Exceptions are some polyphase codes, such as the Frank and P2 codes, which are discrete approximations or variations of FM signals. An additional problem with phase-coded waveforms that can be used in ultrasound is the high peak sidelobes. An exception is the complementary Golay codes, which will be described briefly. Generalized complementary and subcomplementary codes can be potentially useful, if cancellation of sidelobes occurs also away from the delay axis, i.e. for other frequency shifts apart from zero. In this chapter most common sequence families will be presented. Finally Hadamard matrices will be discussed, as they can provide purely orthogonal decoding.

5.1 Phase modulation

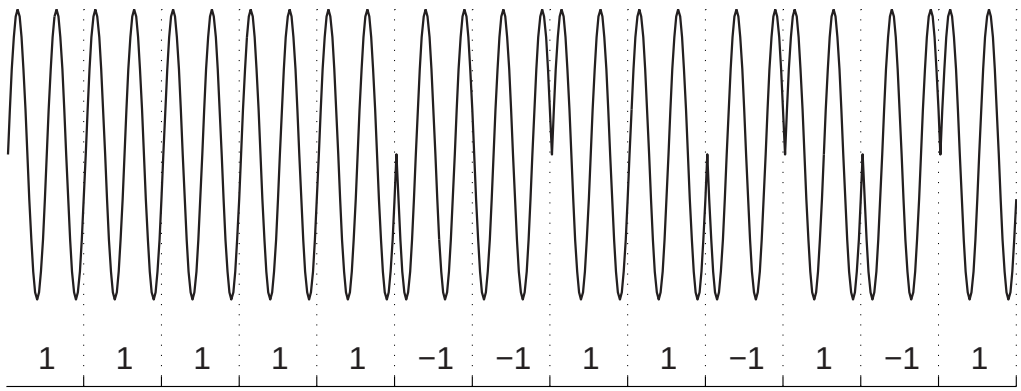

Figure 5.1: Binary-phase coding (Barker-13 code) of a constant-carrier pulse.

The basic idea of phase modulation is a rearrangement of the phases in a signal, so that the cor-

relation integral will cancel out away from the origin. This can be done by taking a long signal of a single carrier pulse (CW), dividing it in N segments of equal length, each consisting of one or more number of cycles (bins), and applying a different initial phase θ_n on each sub-pulse. The number of sub-pulses is the length of the code. For convenience, codes are usually represented by the sequence $\{c_n\}$, where $c_n = \exp(j\theta_n)$. The simplest case of phase coding is the binary or phase-reversal codes, where the phase θ_n of each sub-pulse is either 0° or 180°, and, therefore, c_n consists of 1s and -1s. The phase-coded waveform of the following length-13 code (known as the Barker-13 code):

$$C_{13} = \{1,1,1,1,1,-1,-1,1,1,-1,1,-1,1\}$$

is shown in Fig. 5.1.

Let μ_p be the complex envelope of each sub-pulse. The complex envelope of a phase-coded waveform can be written as:

$$\mu(t) = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} \mu_p(t - nT_p) e^{j\phi_n}, \tag{5.1}$$

where the coefficient $1/\sqrt{N}$ normalizes each sub-pulse with respect to its energy. Using $\mu(t)$ in the definition of the ambiguity function (2.28) gives:

$$\chi(\tau, f_d) = \frac{1}{N} \sum_{n=0}^{N-1} \sum_{m=0}^{N-1} \int_{-\infty}^{\infty} \mu_p(t - nT_p) e^{j\phi_n} \mu_p^*(t - mT_p - \tau) e^{-j\phi_m} dt. \tag{5.2}$$

Introducing a new variable of integration instead of $t - nT_p$, and following the analysis given in [1], the final result is:

$$\chi(\tau, f_d) = \frac{1}{N} \sum_{n=0}^{N-1} \sum_{m=0}^{N-1} e^{j2\pi f_d n T_p} e^{j(\phi_n - \phi_m)} \chi_p\left[(\tau - (n-m)T_p, f_d)\right], \tag{5.3}$$

where $\chi_p(\tau, f_d)$ is the ambiguity function of each single-carrier subpulse. $\chi_p(\tau, f_d)$ is a triangular function along the delay axis in the interval $\pm T_p$ and a *sinc* function on the frequency shift axis, shown in Fig. 2.2. The ambiguity function of a phase-coded waveform is the double summation of the ambiguity function of the sub-pulse, where each component is translated to the position $\tau = (n-m)T_p$ on the delay axis, and is weighted by a phase factor. At the origin there are N ambiguity functions superimposed and superposition is not affected by phase-coding. The summation of the components centered at the origin can be found by setting $n = m$ in (5.3):

$$\chi(\tau, f_d) = \frac{1}{N} \chi_p(\tau, f_d) \sum_{n=0}^{N-1} e^{j2\pi f_d n T_p} = \frac{1}{N} \chi_p(\tau, f_d) \frac{\sin(\pi N T_p f_d)}{\sin(\pi T_p f_d)} \tag{5.4}$$

The peak at the origin is equal to 1. The response at a given delay can be found by collecting the terms in (5.3) centered at the same $\tau = k \cdot T_p$ position. The contribution of these terms is [1]:

$$\chi(\tau, f_d)\,|_{k>0} = \frac{1}{N} e^{-j2\pi f_d k T_p} \sum_{n=0}^{N-1-k} e^{j2\pi f_d n T_p} e^{j(\phi_{n+k} - \phi_n)} \chi_p(\tau - kT_p, f_d) \tag{5.5}$$

In the absence of phase-coding, the summation of the phase factors in (5.5) can be calculated and the component for a given k is:

$$\chi(\tau, f_d)\,|_{k>0} = \frac{N-|k|}{N} e^{j2\pi f_d (N-1+k)T_p} \chi_p(\tau - kT_p, f_d) \frac{\sin[\pi f_d (N-|k|)T_p]}{(N-|k|)\sin(\pi f_d T_p)} \tag{5.6}$$

There are N components superimposed at the origin ($k=0$), and the number of ambiguity functions decreases linearly with $|k|$. When there is no phase-coding, the peak of each k surface is $(N-|k|)/N$, with a maximum of 1 at the origin, and a minimum of $1/N$ at $\tau=(N-1)T_p$, resulting in a triangle auto-correlation function of a long single-carrier pulse. The effect of phase coding is therefore the weighting of each k surface by:

$$\sum_{n=0}^{N-1-k} e^{j2\pi f_d n T_p} e^{j(\phi_{n+k}-\phi_n)} \tag{5.7}$$

5.2 Binary sequences

For binary codes, the complex weighting functions $c_n = \exp(j\theta_n)$ are either 1 or -1, and for $f_d = 0$,(5.7) gives the structure of the auto-correlation function of binary codes:

$$\chi(kT,0) = \sum_{n=0}^{N-1-|k|} c_n c_{n+|k|} \tag{5.8}$$

Applying (5.8) for the Barker code of length 13 shown in Fig. 5.1, the auto-correlation function has been calculated and is shown in the left plot of Fig. 5.2. The compression ratio (the ratio of the compressed pulse and uncompressed pulse widths) is $(NT_p)/T_p$, and is equal to code length N. The sidelobe level is $1/N$ below the main peak.

5.2.1 Barker and other optimal binary codes

The best possible auto-correlation function that can be achieved with phase-reversal codes is the one with uniformly low sidelobes of height $1/N$ shown in Fig. 5.2, where at each bin, there is only one uncompensated component. These are known as the *Barker codes*. These codes would be ideal because of their low sidelobes, if longer lengths were available. Unfortunately, it has been proven that the Barker code of length 13 is the longest available Barker code. A pulse compression system using this code would be limited to a maximum compression ratio of 13 with uniform sidelobe level of $20\log(1/13)$=-22 dB, which is unacceptable for ultrasound imaging.

Optimal binary codes are the sequences whose peak sidelobe of their auto-correlation function is the minimum possible for a given code length. For lengths 2,3,4,5,7,11 and 13, the optimal codes are the Barker codes. For $N>13$, the optimal binary codes have peak sidelobes of $2/N$ or higher.

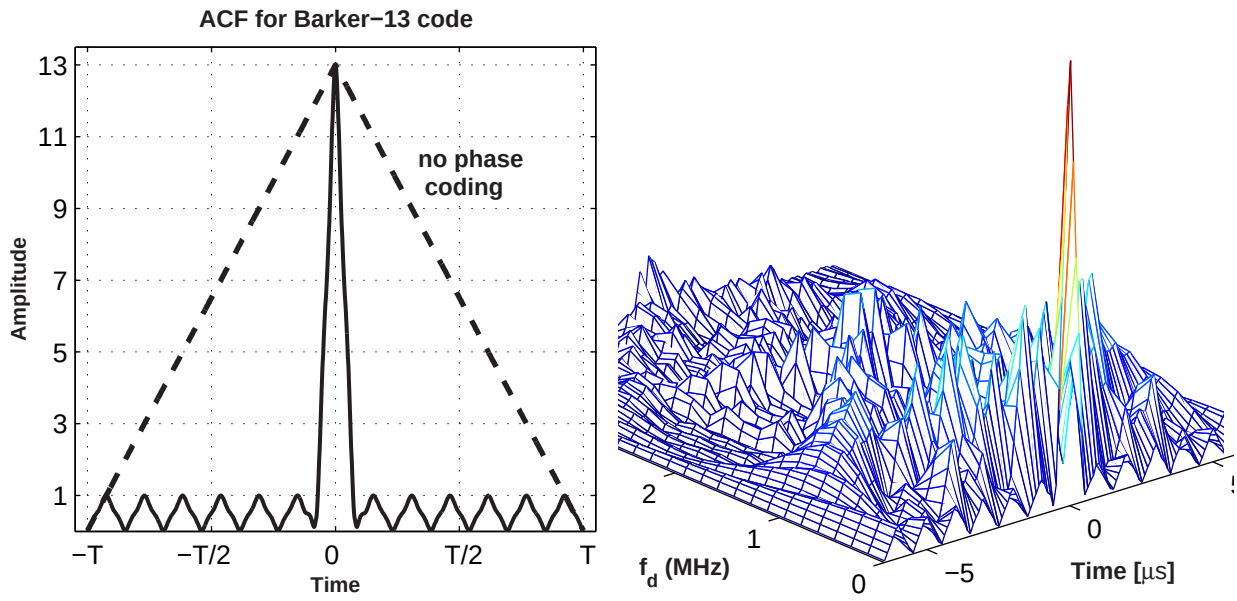

Figure 5.2: Auto-correlation function (left) and ambiguity function (right) of the Barker sequence of length 13 shown in Fig. 5.1.

Computer search is usually employed to find optimal codes. A summary of applicable work [25] indicates little success in finding codes with sidelobes slightly higher than $1/N$ (perhaps $2/N$) but with N so large that the relative sidelobe level is much lower than that for the 13-bit Barker code. A relatively long code with a $2/N$ peak sidelobe level is the following length-28 code [36]:

$$C_{28} = \{1,-1,-1,1,-1,-1,1,-1,-1,1,-1,1,-1,1,-1,-1,-1,-1,-1,-1,1,1,1,-1,-1,1,1,1\}$$

The auto-correlation of this code is shown in Fig. 5.3.

5.2.2 Pseudo-random sequences

There is a large class of sequences which have used extensively in communication, navigation and CW radar systems under the common name *PN (pseudo-noise) sequences*. PN sequences include maximal-length linear feedback shift-register sequences (*m-sequences*), *Gold* and *Kasami* sequences. When such a sequence is repeated periodically, the (periodic) auto-correlation function is repetitive with a period of NT_p and a constant sidelobe level of -1. Thus, in continuous operation, the auto-correlation function has a thumbtack nature, with excellent sidelobe behaviour, which is the reason for the widespread use of such waveforms. However, in order to apply PN sequences in ultrasound, the periodic waveform has to be truncated to one complete sequence. This destroys the low constant sidelobe structure, resulting in peak sidelobes higher than those of the optimal

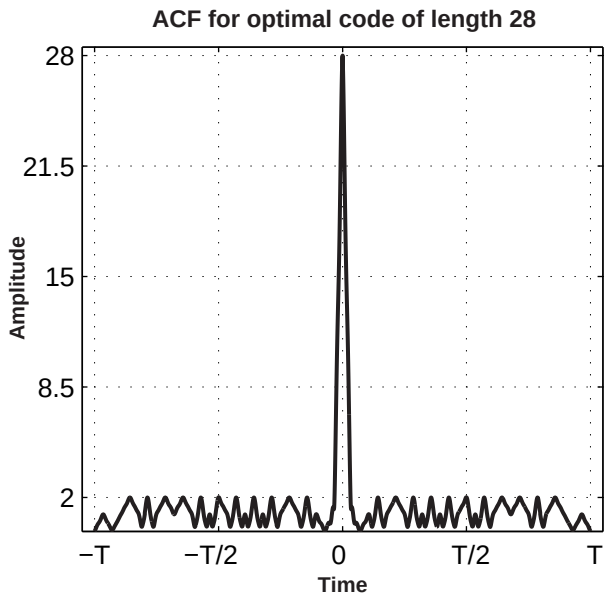

Figure 5.3: Auto-correlation function of the optimal code of length 28. The peak sidelobes have a height of 2, which corresponds to $20\log(2/28) = -23$ dB below the correlation peak.

binary codes for small values of N. The average sidelobe level is -1/2, and for very large N, the standard deviation around this value decreases, resulting in sequences with approximately $1/\sqrt{N}$ sidelobe level. This is still a good sidelobe performance for very large N. An additional advantage of PN sequences is the huge number of sequences in a set with good cross-correlation properties. Unfortunately, both the auto- and the cross-correlation of m-sequences are a function of the code length N, which is a disadvantage for their applicability in medical ultrasound.

One class of periodic sequences which provides larger sets of sequences compared to the m-sequences with good periodic cross-correlation is the class of *Gold sequences*. A set of Gold sequences can be constructed from appropriately selected m-sequences [41]. *Kasami sequences*, derived from Gold sequences, have better cross-correlation properties than the Gold sequences, but a smaller number of signals in the set.

5.2.3 Complementary codes - Golay pairs

One of the main problems associated with all the binary codes discussed so far is the high sidelobe level for short code lengths. Since a delta ambiguity function is utopian, one solution to the problem can be the utilization of a set of waveforms with complex ambiguity functions, which are in some sense as ”different” as possible. By coherently combining the matched-filter responses (the complex ambiguity functions of the set), the restrictions imposed on a single ambiguity function can be bypassed. Guey et.al. [30] proved that for a given energy of the waveforms in a set, the vol-

ume under a combined (or composite) ambiguity surface will always be less than the volume under the ambiguity surface of any single waveform in the set. Furthermore, the minimum occurs when the waveforms is a set of equal-energy orthogonal signals. With this, self-clutter can be reduced without degrading SNR as is the case with the usual mismatch approach.

The simplest application of this method is the utilization of Golay codes, where the combination of ambiguity functions from two Golay codes cancels the sidelobes of the ambiguity function everywhere on the delay axis. Golay codes are pairs of binary codes, belonging to a bigger family of sequences called *complementary pairs*, which consist of two sequences of the same length N whose auto-correlation functions have sidelobes equal in magnitude but opposite in sign. Summing them up results in a composite auto-correlation function with a peak of $2N$ and zero sidelobes. Complementary codes were introduced by Golay in the sixties. Properties and construction methods of Golay codes for a given code length N are discussed in [42] and [43]. Complementary sets, where the auto-correlation functions of more than two sequences have to be added together, have also been derived [44, 45].

In practice, the two sequences of a Golay pair have to be transmitted in different transmit events. Thus, apart from increasing the acquisition rate by a factor of two, some decorrelation on the received rf-data will prevent complete cancellation of the sidelobes. Additionally, the complementarity property of Golay pairs is degraded in presence of ultrasound attenuation, as it will be shown in Chapter 6.

Consideration of extended complementary waveforms, that can cancel the ambiguity everywhere on the $\tau - f_d$ plane has been discussed by Sivaswamy [46]. He showed that from any coded waveform with a duration τ_0, N waveforms of duration $N\tau$ can be constructed, that their auto-correlation sum is zero anywhere except the region $|\tau| < \tau_0$. He named these sequences subcomplementary and he showed that cancellation is total only when the basic waveform is repeated using Hadamard encoding, which will be discussed later in this chapter.

5.3 Polyphase codes

When the phases of each subpulse take more values other than just the 0° and $180^\circ$of binary phases, the resulting waveforms are called *polyphase codes*. The Frank sequences [47] have phases which are the rows of the following matrix:

$$\theta_n = \frac{2\pi}{N}\begin{bmatrix} 0 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 2 & \cdots & N-1 \\ 0 & 2 & 4 & \cdots & 2(N-1) \\ \vdots & & & & \vdots \\ 0 & N-1 & 2(N-1) & \cdots & (N-1)^2 \end{bmatrix} \tag{5.9}$$

The phases of a Frank sequence with length 64 (N=8) and the auto-correlation function of the code are plotted in Fig. 5.4. The auto-correlation function of the periodic Frank code has sidelobes

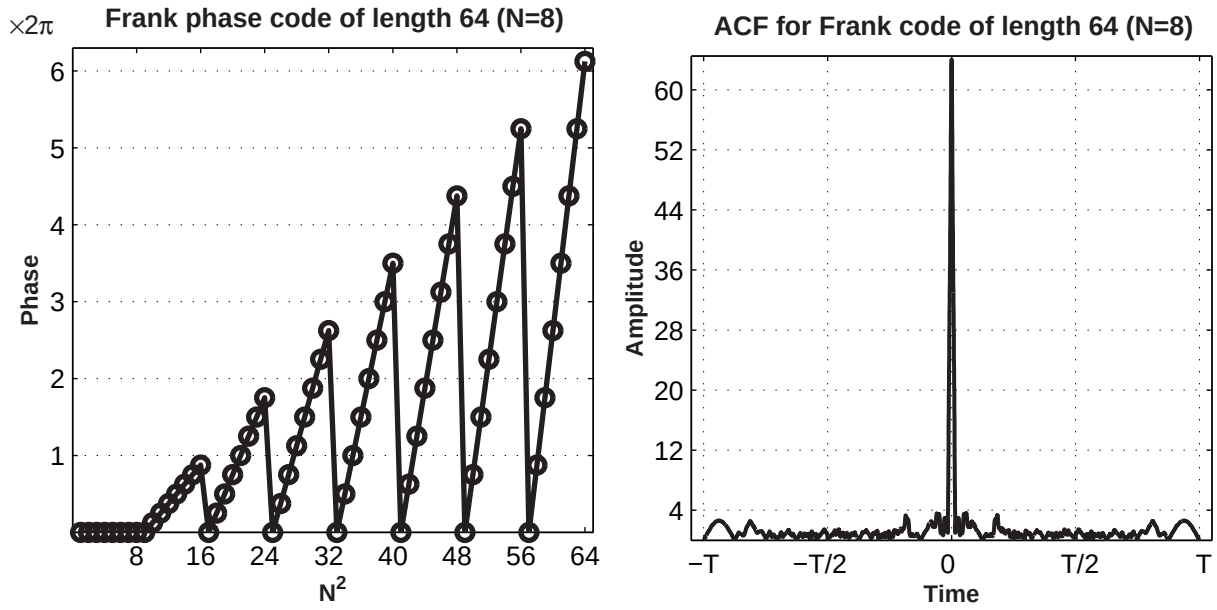

Figure 5.4: Phase coding and auto-correlation function of the Frank sequence with length 64 (N=8). The peak sidelobes have a height of 3.6, which corresponds to $20\log(3.2/64) = -25$ dB below the correlation peak.

of zero [36]. For the aperiodic sequence, the sidelobes decrease as a function of N, and they approach $1/(\pi \cdot N)$ for very large N. In that respect, Frank codes are similar to the PN sequences. Frank polyphase code peak sidelobe ratio is superior to that of most binaray codes found from computer searches for high values of N. For instance, a 2500-element Frank code (N=50) has a peak sidelobe ratio of -45 dB, without any amplitude weighting. In contrast, the peak sidelobe ratio of the compressed chirp waveform is only -13.2 dB independent of the time-bandwidth product.

The ambiguity response resembles the ridgelike characteristics of the linear FM signal, in contrast to the thumbtack ambiguity of PN sequences. This makes the Frank codes much tolerable to frequency shifts than the binary codes. However, the N-dependence in the sidelobe behaviour of the Frank code is a disadvantage for ultrasound, where N is limited. For example, the shortest Frank code having one cycle per bin at a carrier frequency of 4 MHz, for N=8 would be 16 μs long, with peak sidelobes of -25 dB. This is still insufficient for high resolution ultrasound imaging.

The spectrum of the Frank coded waveform contains higher harmonics and sub-harmonics of the carrier frequency (Fig. 5.5), and therefore some degradation in sidelobe performance is expected, when a band-limiting ultrasound transducer is excited with this signal.

Lewis and Kretschmer [48] have rearranged the phases of the Frank code in order to achieve higher

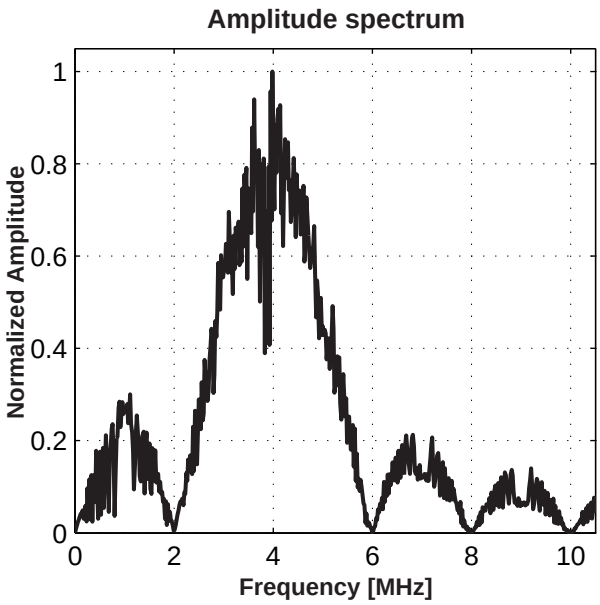

Figure 5.5: Amplitude spectrum of the Frank code.

tolerance of receiver bandlimiting before pulse compression. They refer to the modified sequences as $P1$ and $P2$ codes.

Other polyphase waveform sets with similar auto-correlation functions and good cross-correlation properties, such as the Zadoff-Chu, P4 and generalized chirp-like (GCL) sequences have been discussed in the literature [49]. Discussion of matched filtering of polyphase codes can be found in [50] and [51].

5.4 Hadamard matrices

The Hadamard matrix is a $N \times N$ square matrix consisting of 1 and -1, which is orthogonal both along its rows and along its columns. If $H_{m,n}$ are the entries of the Hadamard matrix, the orthogonality property is described mathematically as:

$$\begin{aligned} \sum_{n=1}^{N} H_{m,n} \cdot H_{m',n} &= 0, \quad m = 1 \ldots N, \quad m' = 1 \ldots N \quad (m \neq m') \\ \sum_{m=1}^{N} H_{m,n} \cdot H_{m,n'} &= 0, \quad n = 1 \ldots N, \quad n' = 1 \ldots N \quad (n \neq n') \end{aligned} \tag{5.10}$$

The smallest Hadamard matrix is the 2×2 matrix given by:

$$\mathbf{H}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}. \tag{5.11}$$

The $2n \times 2n$ Hadamard matrix of length $2N$ can be constructed recursively by the length $N = n$ $n \times n$ Hadamard matrix as following:

$$\mathbf{H}_{2N} = \frac{1}{\sqrt{2}} \left[\begin{array}{c|c} \mathbf{H}_N & \mathbf{H}_N \\ \hline \mathbf{H}_N & -\mathbf{H}_N \end{array} \right] \tag{5.12}$$

The rows of the Hadamard matrix of (5.12) are the so called Walsh functions, and (5.12) represents the kernel for decomposing a function in a set of orthogonal Walsh functions.[1]

An important property of the Hadamard matrices is that the inverse of the normalized Hadamard matrix is the transpose matrix:

$$\mathbf{H}^{-1} = \mathbf{H}^{\mathrm{T}} \quad , \quad \mathbf{H}\mathbf{H}^{\mathrm{T}} = \mathbf{I} \tag{5.13}$$

For example, let two signals $\mathbf{X}^{\mathrm{T}} = [\mathrm{x}_1 \mathrm{x}_2]$ be encoded by the 2×2 Hadamard matrix:

$$\mathbf{H}\mathbf{X} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} x_1 + x_2 \\ x_1 - x_2 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \tag{5.14}$$

The two signals can be restored by applying the transpose Hadamard matrix on the encoded signal set:

$$\mathbf{H}^{\mathbf{T}}\mathbf{Y} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} y_1 + y_2 \\ y_1 - y_2 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} (x_1 + x_2) + (x_1 - x_2) \\ (x_1 + x_2) - (x_1 - x_2) \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \tag{5.15}$$

If the rows of the Hadamard matrix correspond to channel coding and the columns correspond to transmit events, the above example can separate two signals x_1, x_2 transmitted simultaneously by different channels, when two emissions are used. Such Hadamard-encoding of ultrasonic transmission can find application in synthetic aperture techniques, which will be discussed in Chapter 9.

5.5 Sidelobe reduction for phase-encoded sequences

As opposed to the sidelobes of the linear FM signal, which generally decrease as a function of time displacement away from the peak, the sidelobes of most phase-coded signals tend to have an almost constant level extending out to $\pm T$ seconds on either side of the compressed pulse. The long time duration of the range sidelobes of the phase-coded signals will cause more serious problems in masking weaker signals that the range sidelobes of the FM type of signal, particularly in a dense medium [25].

[1] The name Walsh-Hadamard matrix has also been used in the literature.

Mismatched filters for sidelobe reduction in FM type of signals are designed easily by weighting the spectrum with appropriate window functions. In phase-coded signals, the lack of symmetry in the signal spectrum requires optimization methods. The filter coefficients are calculated by minimization of quantities such as the integrated sidelobe level (ISL filters) or the peak sidelobe level (PSL filters). Mismatched filter design techniques can be found in [51] and [52]. Since such mismatched filters are not optimized for signal-to-noise ratio, there is a significant degradation in SNR gain.

5.6 Disadvantages of phase-coding for ultrasound imaging

There is a number of reasons why phase codes can only have limited use in ultrasound imaging.

- Phase codes are based on abrupt changes in phase and their spectrum contains higher harmonics that the ultrasonic transducer smooths out. The effect of this bandwidth reduction in the actual transmitted codes is that sidelobe levels are higher than the expected $1/N$ or $2/N$ levels. Additionally, the orthogonality property among the signals in the set degrades. Therefore, pulse compression with phase codes applied in ultrasound systems has a worse performance than what has been presented in this chapter, due to the limited bandwidth.

- The sidelobes of the linear FM signal are relatively independent of the time-bandwidth product, once the spectrum shape is fixed. Practically, FM signals with a TB greater than 25 have similar sidelobe behaviour and, thus, the same sidelobe reduction techniques are applicable and equally effective regardless of the TB. In contrast, the level of the range sidelobes in phase-coded signals is a function of the time-bandwidth product, i.e. the code length. In radar systems, where phase codes have been used successfully, code lengths as high as 1,000 are possible, which result in pulse compression systems with range sidelobes down to -45 to -50 dB without any weighting. In ultrasound, the length of the code cannot exceed 80-100 at most, which in conjuction with the band-limiting effect of the transducer will result in very poor sidelobe performance of about -25 dB below the auto-correlation peak.

- Most binary phase codes share the undesirable property of being quite sensitive to frequency shifts. The ambiguity function of such signals will generally have a central spike on a wide pedestal of range-frequency sidelobes. A search for binary codes with good autocorrelation functions is quite useless in ultrasound, if the ambiguity function away from the delay axis is unsuitable [1]. The Barker or the Golay codes, for instance, do not retain their good compression output even with moderate frequency shifts (Fig. 5.2b).

CHAPTER

SIX

Ultrasound imaging with coded excitation-Simulation results

6.1 Intensity considerations

The acoustic intensity of a wave is the average flow of energy through a unit area normal to the propagation direction in unit time [6]. The intensity is the average of the rate of work done per unit area by one element of fluid on an adjacent element. That is [1]:

$$I =< pu >_t = \frac{1}{T}\int_0^T pu\,dt, \tag{6.1}$$

the integration taken over one cycle of the harmonic field. For a plane, progressing wave:

$$\begin{aligned} p(t) &= p_0 \exp(j(\omega_0 t - kz)) \\ u(t) &= \frac{p(t)}{Z}, \end{aligned} \tag{6.2}$$

where ω_0 is angular frequency, $k = \omega_0/c$ wave number, and z is the distance traveled in the direction of propagation. So

$$I =< pu >_t = \frac{1}{2\pi/\omega_0}\int_0^{\frac{2\pi}{\omega_0}} \frac{p_0^2}{Z}\exp(2j(\omega_0 t - kz))dt = \frac{p_0^2}{2Z}. \tag{6.3}$$

The SI unit of intensity is W/m^2, but a more commonly used unit in medical ultrasound is mW/cm^2.

The measure of intensity given by (6.3) is only meaningful for a continuous wave. Since the ultrasound fields generally used in medical practice are pulsed and strongly varying with regard to spatial position, it is necessary to specify how the intensity is to be calculated. The instantaneous intensity is given by

$$I_i(t,\vec{r}) = \frac{p^2(t,\vec{r})}{\rho c}. \tag{6.4}$$

[1] The intensity definitions in this section are taken from Jensen [2] after permission.

Table 6.1: Highest known acoustic field emissions for commercial scanners as stated by the United States FDA (The use marked (a) also includes intensities for abdominal, intra-operative, pediatric, and small organ (breast, thyroid, testes, neonatal cephalic, and adult cephalic) scanning). Table is reproduced from Jensen [2] after permission.

	I_{spta} mW/cm^2		I_{sppa} W/cm^2		I_m W/cm^2	
Use	In Situ	Water	In Situ	Water	In Situ	Water
Cardiac	430	730	65	240	160	600
Peripheral vessel	720	1500	65	240	160	600
Ophthalmic	17	68	28	110	50	200
Fetal imaging (a)	46	170	65	240	160	600

The field can be characterized by the spatial and temporal peak of I_i, denoted I_{sptp}, or it can be averaged over time giving the spatial peak temporal averaged intensity

$$I_{spta} = \frac{1}{T'} \int_0^{T'} I_i(t, \vec{r}_{max}) dt, \tag{6.5}$$

where T' is the period from pulse to pulse and $\vec{r}_{max}$ denotes the position of maximum intensity. Other intensity measures are used which include spatial and/or temporal averaging. I_{sppa} is the spatial peak pulse average intensity, which is calculated as the mean intensity over the pulse duration at the position of maximum intensity in the field. It is given by [53]:

$$I_{sppa} = \frac{\int p^2(t) dt}{\rho c \cdot PD}. \tag{6.6}$$

PD is the pulse duration and is 1.25 times the interval between the time when the time integral of intensity for an acoustic pulse at a point reaches 10%, and when it reaches 90% of the pulse intensity integral given by

$$P_{ii}(t, \vec{r}) = \int_{-\infty}^{t} I_i(t', \vec{r}) dt'. \tag{6.7}$$

All commercial scanners intended for clinical use have to comply with intensity limits, which are set by the United States Food and Drug Administration (FDA). The currently reported values are given for both in situ and water measurements, and are shown in Table 6.1 according to the FDA [54]. High peak intensity acoustic pressure waves can generate acoustic cavitation, i.e. localized rapid generation and collapse of gas-filled bubbles. The collapse is adiabatic and can cause tremendous localized temperature rise and tissue damage. Another effect of high-intensity ultrasound waves is the non-linear propagation effects, which is the reason why most ultrasound scanners use driving voltages that generate peak intensities well below the FDA limits. For instance, for the transducer used, the intensities measured in water by the company (B-K Medical) with a conventional pulse of 70V were I_{spta}=3.19 mW/cm^2 and I_{sppa}=66.4 W/cm^2 at the focal point. For comparison, the FDA limits for adult imaging of peripheral vessels [55] are I_{spta}=1500 mW/cm^2

(spatial-peak-temporal-average) and I_{sppa}=240 W/cm^2 (spatial-peak-pulse-average) in water. Both intensities are well below the limits.

I_{sppa} is the mean intensity over the pulse duration at the focal point and therefore does not change significantly, when the duration of the signal changes. Therefore, coded excitation will not affect it. However, the I_{spta} is an integration over time and is increased when the duration of the pulse is increased. For a chirp of duration 20 μs, the I_{spta} is about 60 times higher depending on the applied code, i.e. 196 mW/cm^2, which is still more than 7 times lower than the specified safety limits.

Recommendations have been put forward by the American Institute of Ultrasound in Medicine [55] on the maximum ultrasound energy allowed to be radiated into human tissue for diagnostic purposes. The recommendation was made in 1988 and was based on studies of tissue damage due to ultrasound exposure. The safety standard reads:

> In the low megahertz range there have been (as of this date) no independently confirmed significant biological effects in mammalian tissues exposed *in vivo* to unfocused ultrasound with intensities[a] below 100 mW/cm^2, or to focused ultrasound[b] with intensities below 1 W/cm^2. Furthermore, for exposure times[c] greater than 1 second and less than 500 seconds (for unfocused ultrasound) or 50 seconds (for focused ultrasound) such effects have not been demonstrated even at higher intensities, when the product of intensity and exposure time is less than 50 joules/cm^2.
>
> [a] Spatial-peak temporal average intensity.
> [b] Quarter power beam width smaller than four wavelengths or 4 mm, whichever is less at the exposure frequency.
> [c] Exposure time includes off-time as well as on-time for repeated pulse regimes.

6.2 Expected signal-to-noise ratio improvement

The anticipated improvement in signal-to-noise ratio by using coded excitation in ultrasound imaging, has been evaluated with the simulating program Field II [56]. A single-element transducer focused at 10 cm, with diameter of 1.57 cm, a nominal frequency of 4 MHz and a 65% -6 dB bandwidth has been simulated. The transducer parameters correspond to those of the mechanically-rotating single-element probe that has been used in the experiments, which will be presented in the next chapter. From hydrophone measurements in a water tank, peak intensities were calculated for pulsed and coded excitation. It was found that pulse and coded excitation of equal peak voltage amplitude resulted in a measured I_{sptp} (spatial-peak-temporal-peak) about 1.6 times greater for the

coded signal at the acoustical focus. This is due to different focusing properties of the transducer for a much longer signal. The relationship is non-linear and is between 1.1 and 1.8 depending on the applied voltage amplitude. For instance, a conventional pulse of 80 V and a chirp signal of 50 V give the same I_{sptp} at the focal point. The simulations showed a similar dependence, and therefore for all SNR gain simulation results, the excitation voltages were normalized to yield the same I_{sptp} at the focal depth for all excitation signals. The calculation of the SNR has been based on the following equation:

$$SNR_i = \frac{\int_0^{T_P} p_i^2(t,\vec{r}_i)dt}{P_{noise}}, \tag{6.8}$$

where p is the pressure field at the axial position $\vec{r}_i$, T_P is the total duration of the received echoes and P_{noise} is the power of the simulated band-pass thermal (Johnson) noise of the system. The energy of the received echoes has been calculated as an integration over time of the squared pressure field. The results have been normalized so that the peak SNR in case of pulsed excitation is always 60 dB.

Fig. 6.1 shows simulation results on the expected SNR gain from coded excitation for four different coded signals, when there is no attenuation in the medium. The four coded signals (which are shown in Fig. 6.2), are a standard linear FM signal, the amplitude-tapered FM signal presented in Chapter 4, a non-linear FM signal and a Golay sequence of length 40. The right plots of Fig. 6.2 show the actual transmitted coded signals after convolution with the transducer impulse response. The energy of the signals is reduced due to the transducer filtering. The energy reduction of the transmitted coded signals is 74% for the linear FM, 72% for the tapered linear FM, 41% for the non-linear FM and 23% for the Golay code. The reduced transmitted energy of the linear FM signals explains the results of Fig. 6.1. The expected SNR improvement is 11.1 dB for the linear FM signal, 10.4 dB for the tapered linear FM signal, 14.6 dB for the non-linear FM signal and 16 dB for the Golay code. The SNR gain is independent of depth and the overall SNR reduction is due to the $1/r$ dependence of acoustic diffraction.

Fig. 6.3 shows simulation results on the expected SNR gain of the same coded signals when attenuation of 0.5 dB/[MHz$\times$cm] has been simulated for the medium. The expected SNR improvement is depth-dependent and it improves significantly for the linear FM signals. At the acoustic focus is 10.9 dB for the linear FM signal, 10.2 dB for the tapered linear FM signal, 12.8 dB for the non-linear FM signal and 14.5 dB for the Golay code. At higher depths where we expect to benefit from coded excitation the most, the expected SNR gain of the linear chirps relative to the pulsed excitation increases. At a depth of 20 cm, the SNR for the pulsed excitation is reduced by 5 dB, due to attenuation, the SNR for the Golay code is reduced by more than 6 dB, while the SNR of the linear FM excitation remains constant at 52 dB at 20 cm. Thus, the SNR gain for the linear FMs is around 17 dB, while it is around 14 dB for the non-linear FM and Golay codes.

In the next set of simulation results shown in Fig. 6.4, matched filtering has been applied on the received rf-data and the SNR has been calculated from the energy of the compressed pulses.

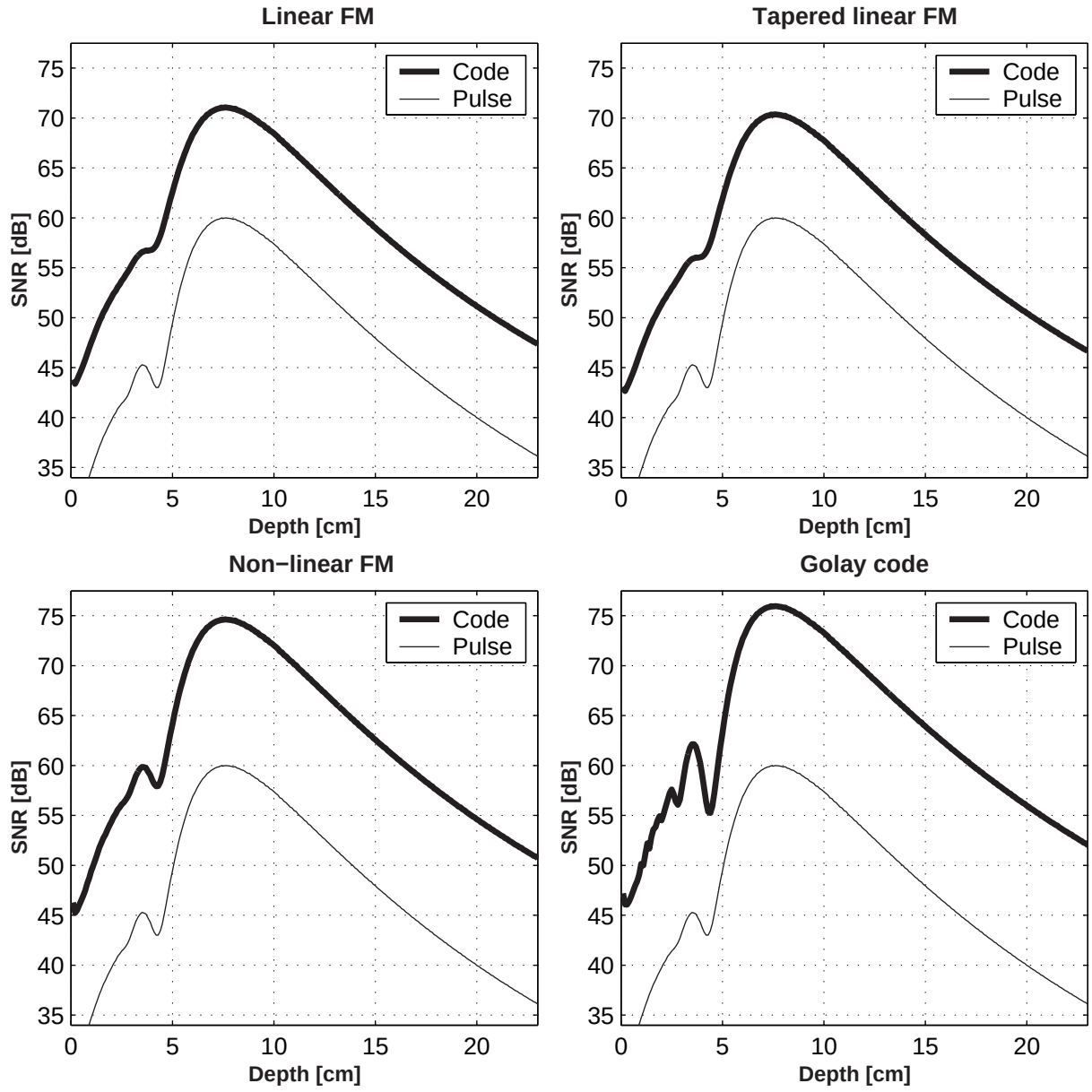

Figure 6.1: Simulation results on the expected SNR improvement from coded excitation for four different coded signals using Field II. There is no ultrasonic attenuation in the simulated medium. The higher transmitted energy of Golay codes results in higher SNR gain compared to the linear FM signals.

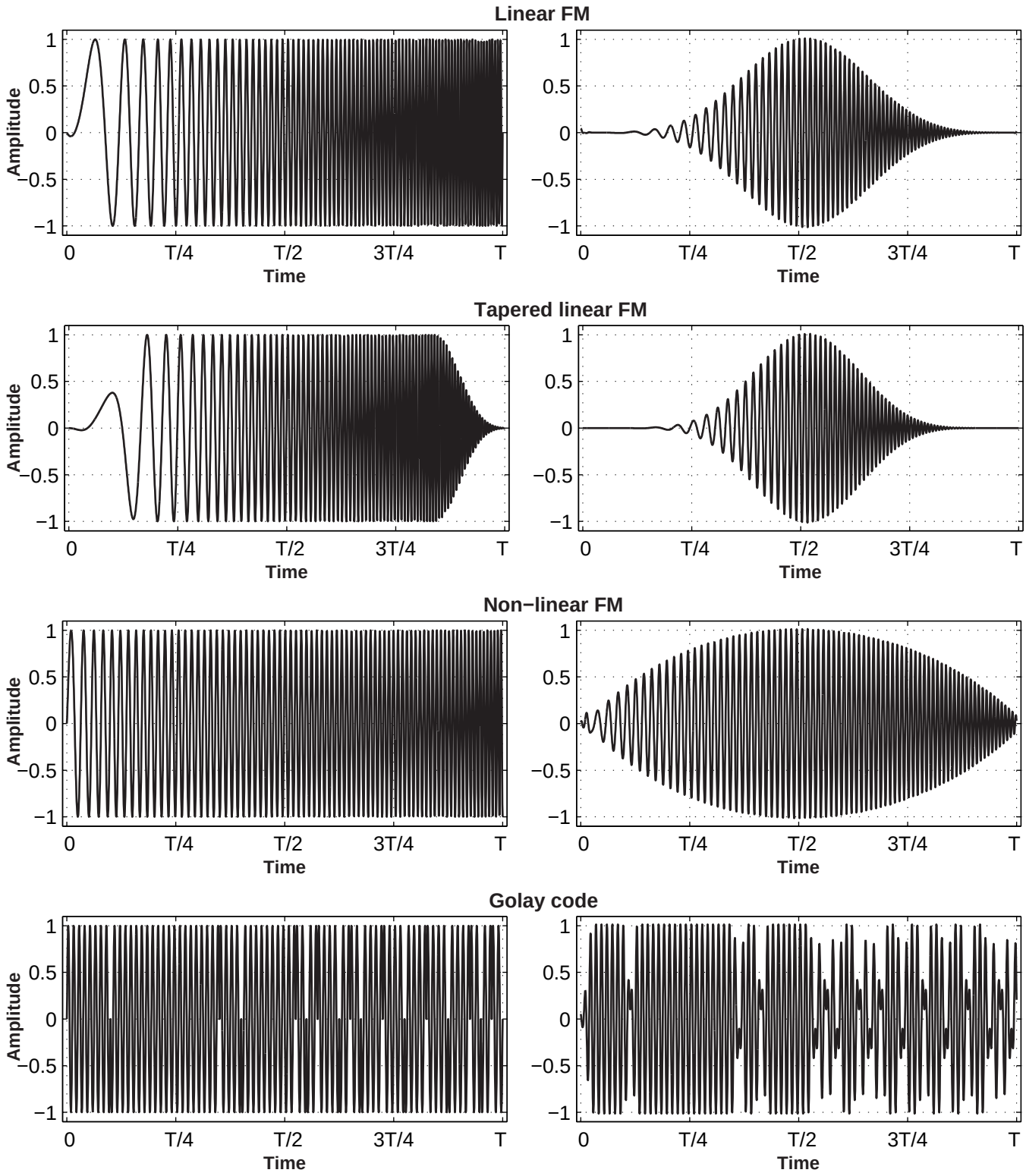

Figure 6.2: The four coded excitation signals used in the SNR simulations. On the right plots are the actual propagating signals after convolution with the transducer impulse response. The presence of the transducer affects the transmitted energy of the linear FM signals the most.

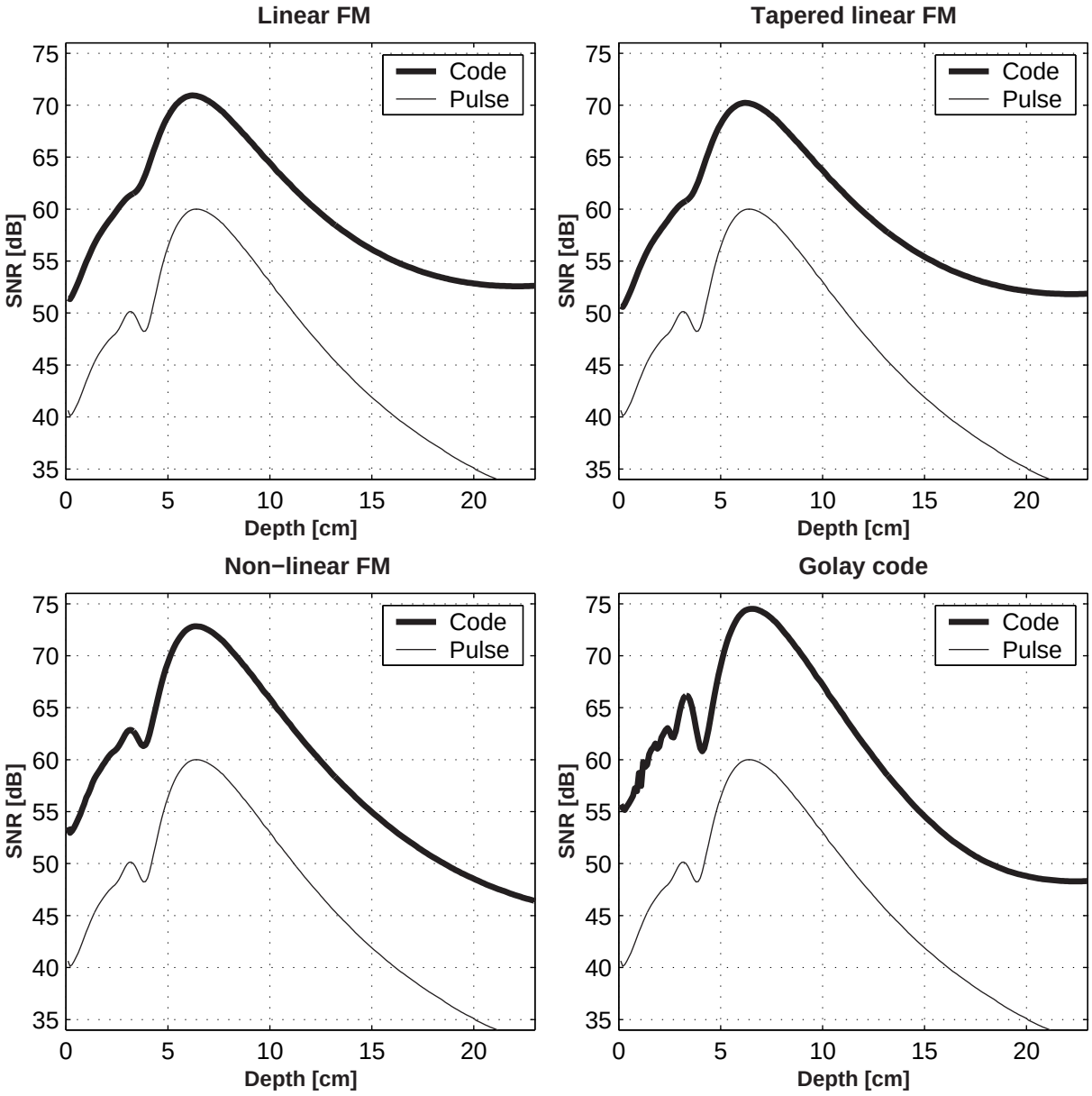

Figure 6.3: Expected SNR improvement for various codes in tissues with attenuation of 0.5 dB/[MHz×cm]. The linear FM signals exhibit higher SNR gain relative to the pulsed excitation than the non-linear FM and Golay-coded signals.

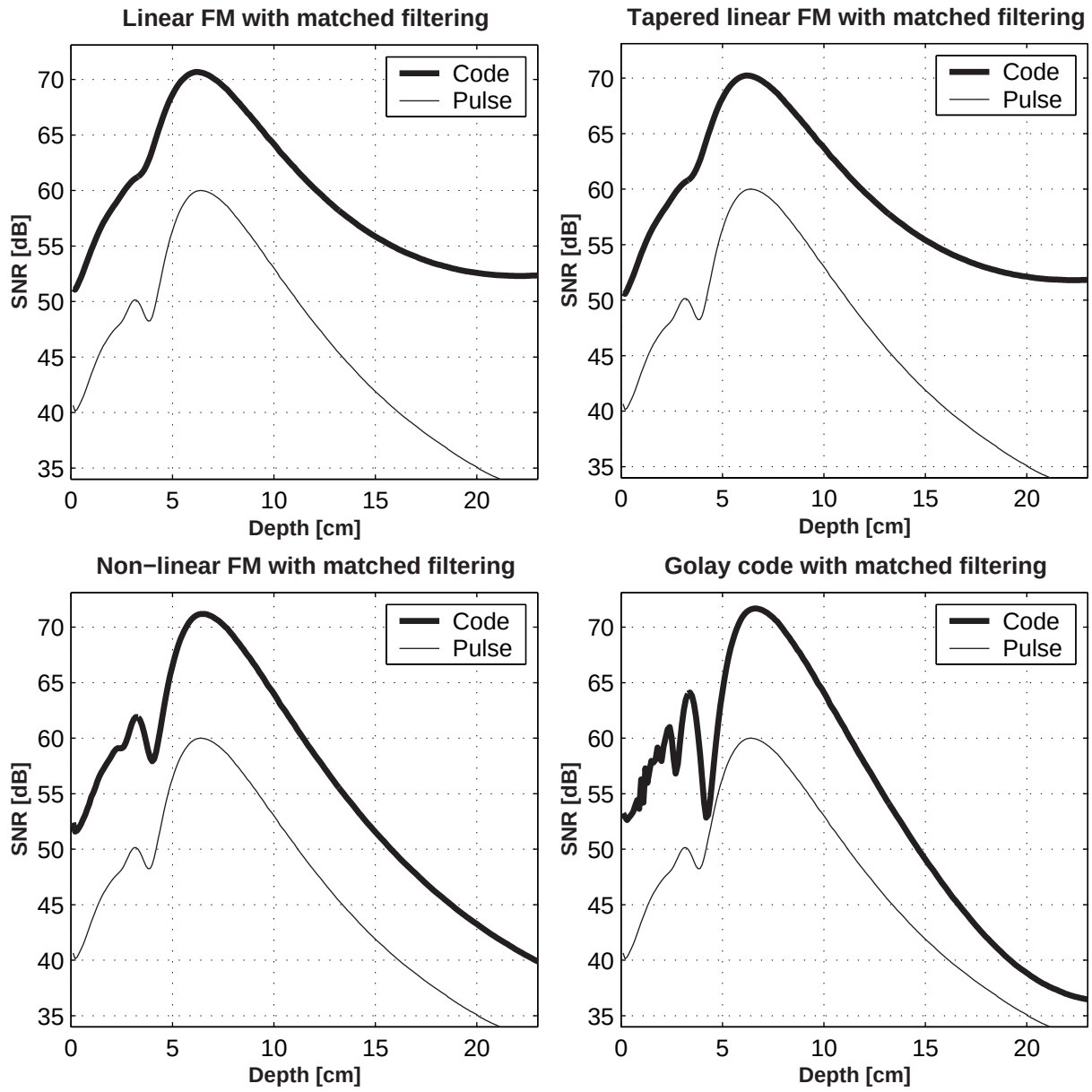

Figure 6.4: Expected SNR improvement in tissues with attenuation of 0.5 dB/[MHz×cm] after matched filtering.

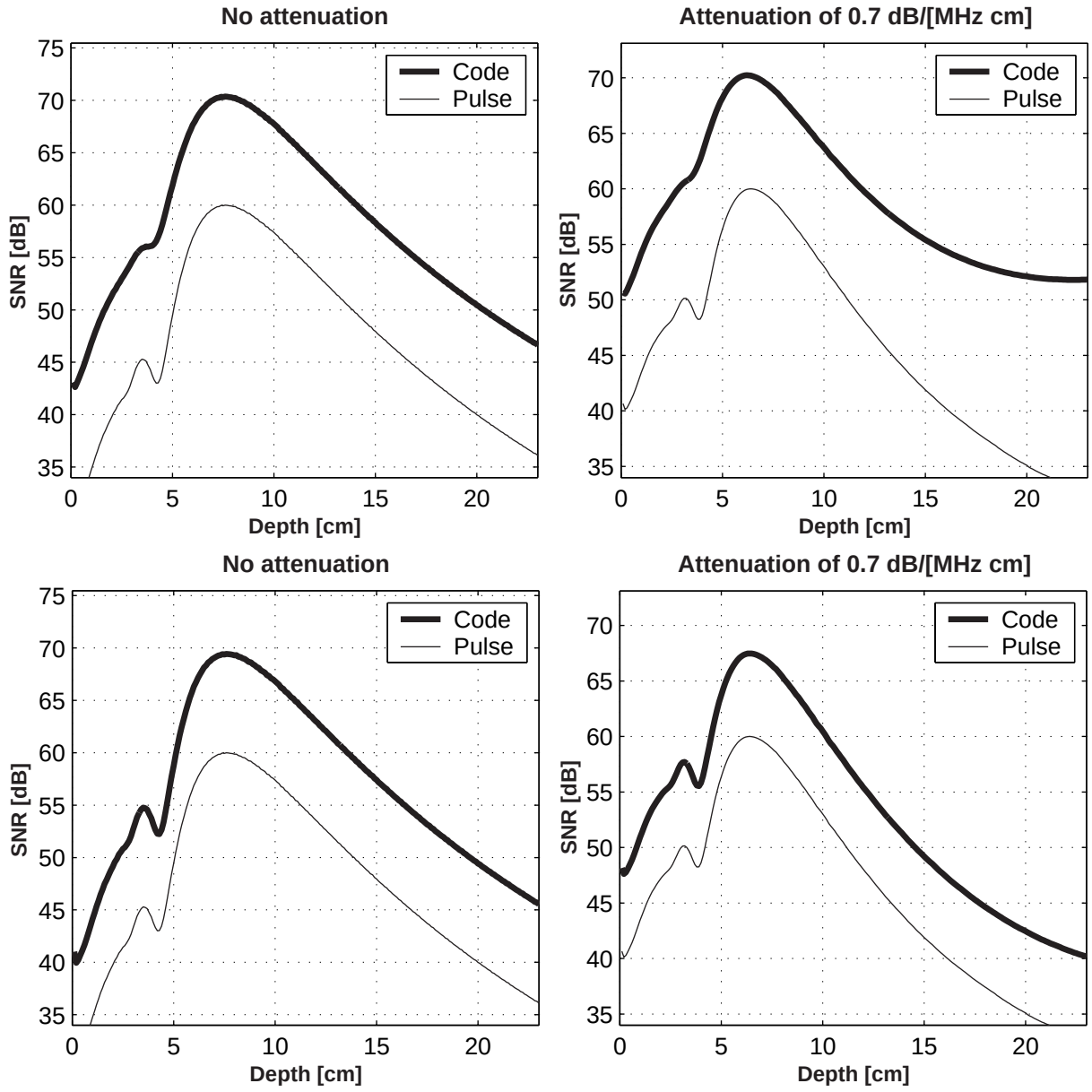

Figure 6.5: The effect of mismatched filtering for the tapered linear FM signal on the expected SNR improvement in tissues without attenuation (left) and with attenuation of 0.5 dB/[MHz×cm] (right). The upper two plots show SNR gain for pure matched filtering and the lower plots show SNR gain for mismatched filtering.

Matched filters have been applied on all signals. The theoretical compression output (without the effect of the attenuation) of the first linear FM signal in the presence of the transducer is the gray line of Fig. 4.2b, exhibiting distant sidelobes of -35 dB. The matched filter response of the second tapered linear FM signal has sidelobes of -68 dB. The matched filter response of the third non-linear FM signal is the one shown in Fig. 3.6 with sidelobes of -36 dB. As it will be shown later in this chapter, the sidelobe levels in the compression output of the last two signals will be higher in the presence of attenuation.

The graphs in Fig. 6.4 show the effect of distortions from frequency shifting due to attenuation on the SNR loss. The SNR improves with depth only for the linear FM signals, while it is reduced for the non-linear FM and Golay signals. These graphs show the robustness of linear FM signals in attenuation with respect to SNR. Further in this chapter, the robustness of different coded signals in attenuation with respect to sidelobes will be investigated.

The last set of simulation results investigates the effect of weighting of the filter amplitude spectrum (mismatched filtering) on the expected SNR gain, relative to the pure matched filtering. Comparison of the two left plots in Fig. 6.5 shows that when the medium has no attenuation, mismatched filtering causes the expected from theory 1 to 2 dB SNR gain loss. Unfortunately, in the presence of attenuation (right two plots), the SNR gain loss caused by mismatching is severe. The upper right plot of Fig. 6.5 is the same as the upper right plot of Fig. 6.4 and is repeated here for convenience. For instance, at depth of 20 cm, there is a SNR gain loss of 9 dB, from matched to mismatched filtering. This will result in a SNR gain diminishing from 17 dB when matched filtering is applied to only 8 dB for mismatched filtering. This result has not been considered in previous literature for the application of coded signals in medical ultrasound imaging. The significance of this is that, **due to attenuation, very low sidelobes attained by mismatching and high SNR gain are incompatible for ultrasound imaging**. If a matched filter is used, a SNR gain of almost 17 dB at 20 cm is theoretically possible, while the sidelobes level will entirely depend on the transducer weighting. On the other hand, if a mismatched filter is used, sidelobes can be kept very low and independent on the actual transducer impulse response and attenuation, but the SNR gain will be less than 10 dB.

6.3 Imaging with linear FM signals- Simulation results using Field II

Fig. 6.6 and Fig. 6.7 show simulated images using conventional pulsed and linear FM excitation. The simulation parameters correspond to the single-element mechanically-rotating transducer with a nominal frequency of 4 MHz and a 65% -6 dB that has been used in the experiments. Eight point scatterers along the axial direction are imaged. The pulsed excitation is a sinusoidal signal of 4 cycles with Hanning apodization, which matches the transducer impulse response. The FM-coded

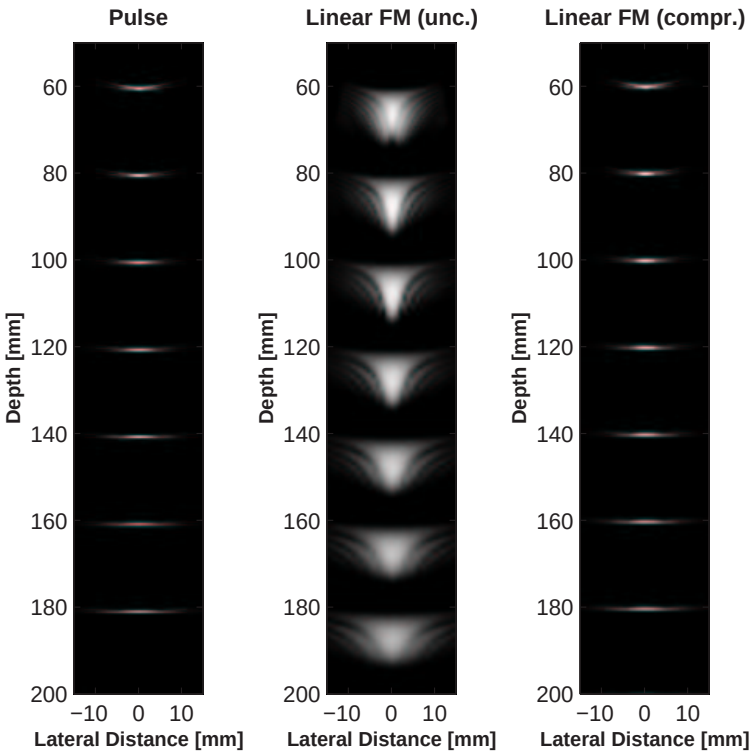

Figure 6.6: Pulsed vs. FM-coded excitation imaging in a medium with no attenuation.

excitation is the one whose mismatched filter response is shown in Fig. 4.6b. The simulated images shown in Fig. 6.6 and Fig. 6.7 correspond to a non-attenuating medium and a medium with attenuation of 0.7 dB/[MHz$\times$cm] respectively. The visual appearance of pulsed and coded images is very similar. Quantitative evaluation of image resolution will be given in a following section. The sidelobe performance of pulse compression can be seen from Fig. 6.8, where the compressed central rf-lines are plotted. With no attenuation, the sidelobe levels are very close to what is expected from Fig. 4.6b, i.e. the sidelobes are close to -88 dB for all depths. In the attenuating medium, the sidelobes increase with depth, but remain below -60 dB for all depths. Fig. 6.9 shows the central rf-lines for tapered linear FM excitation with mismatched filtering when the actual transducer impulse response has been used in the simulations. Compression is robust to both attenuation and transducer weighting. A plot of the central rf-lines for tapered linear FM excitation with pure matched filtering is shown in Fig. 6.10. In this case, compression is good in a non-attenuating medium, but it degrades in the presence of attenuation. This excitation/compression scheme yields the highest SNR gain of 17 dB at 20 cm, according to the simulation results presented in the last section. It can be used if SNR gain is the critical imaging parameter, and the resulting resolution and sidelobe levels shown in Fig. 6.10 are acceptable for a given application.

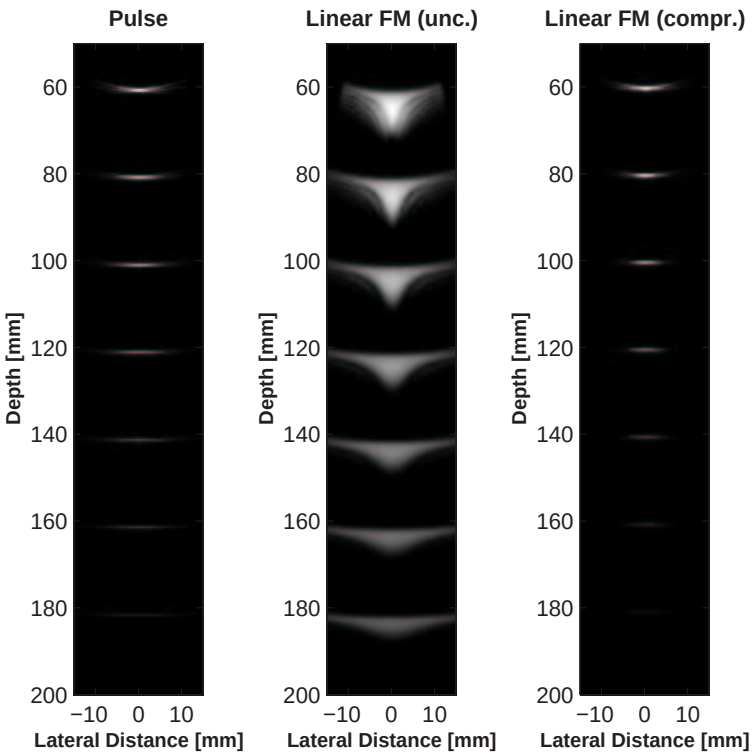

Figure 6.7: Pulsed vs. FM-coded excitation imaging in a medium with attenuation of 0.7 dB/[MHz×cm]

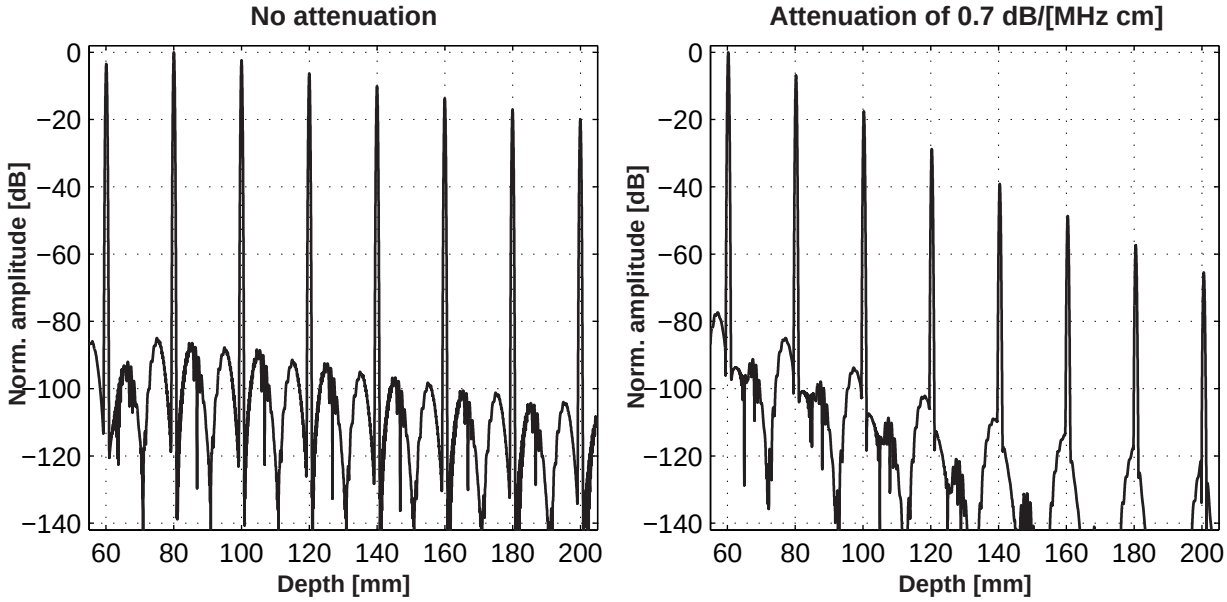

Figure 6.8: The minimal effect of attenuation in sidelobe levels using tapered linear FM excitation with mismatched filtering. The graphs show the central rf-lines of the coded images in the absence (left) and presence (right) of attenuation in the medium.

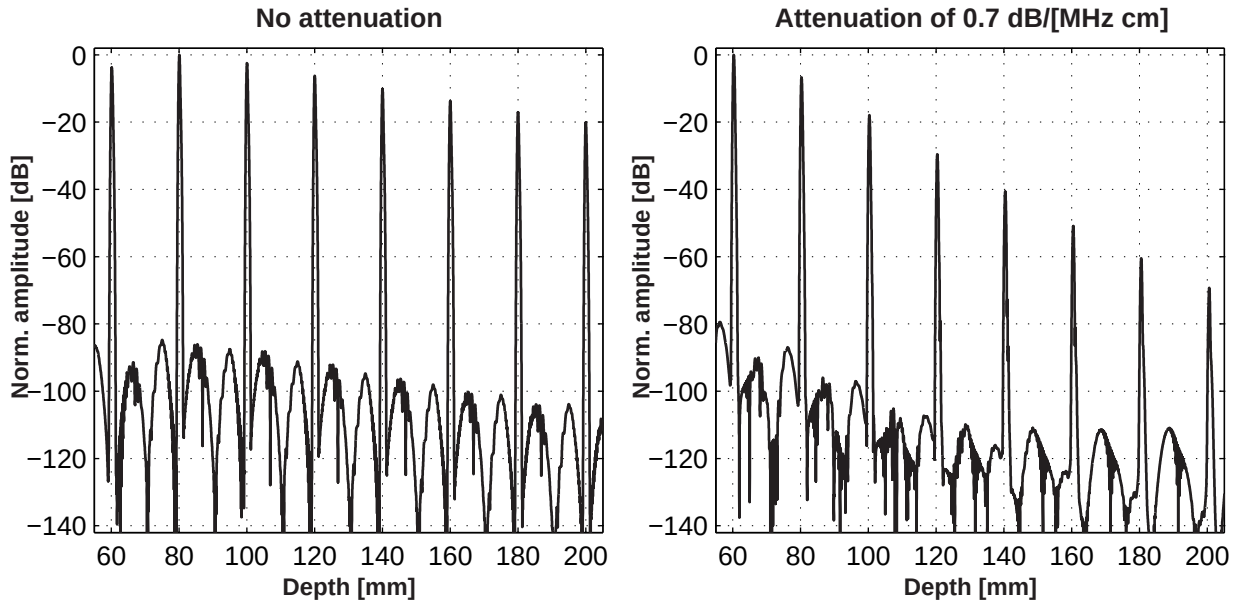

Figure 6.9: The same as in Fig. 6.8 (tapered linear FM excitation with mismatched filtering) but with the measured transducer impulse response used in the simulations. The effect of the actual transducer impulse response in sidelobe levels is small.

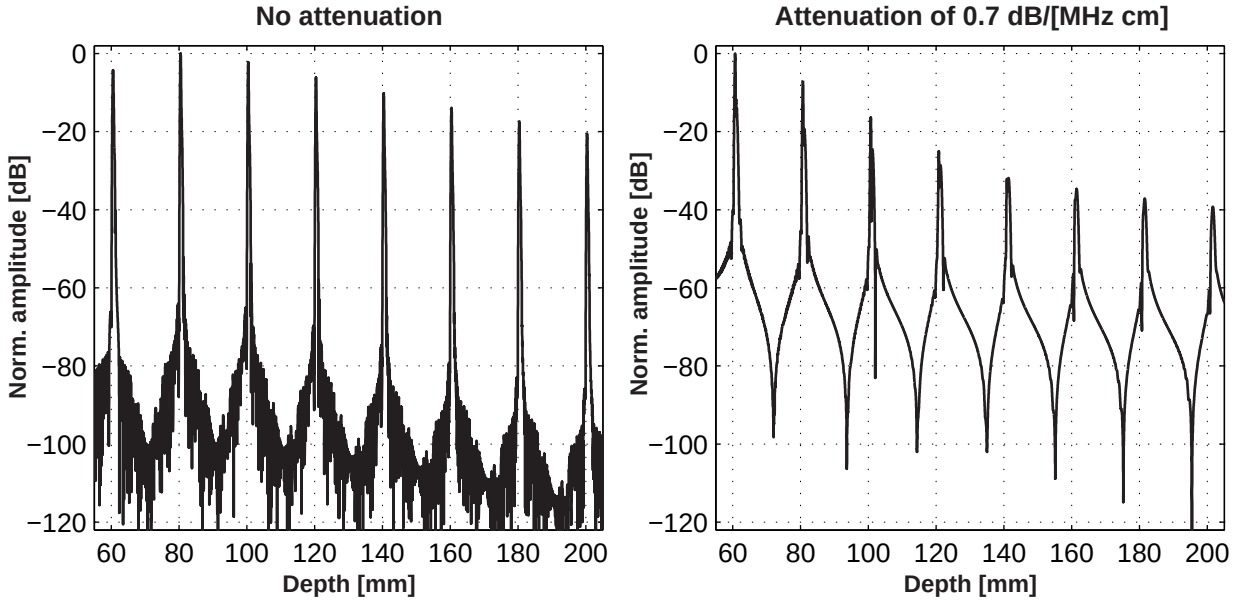

Figure 6.10: The effect of attenuation and transducer weighting for tapered linear FM excitation with pure matched filtering. In the presence of attenuation, compression is very sensitive to the transducer impulse response.

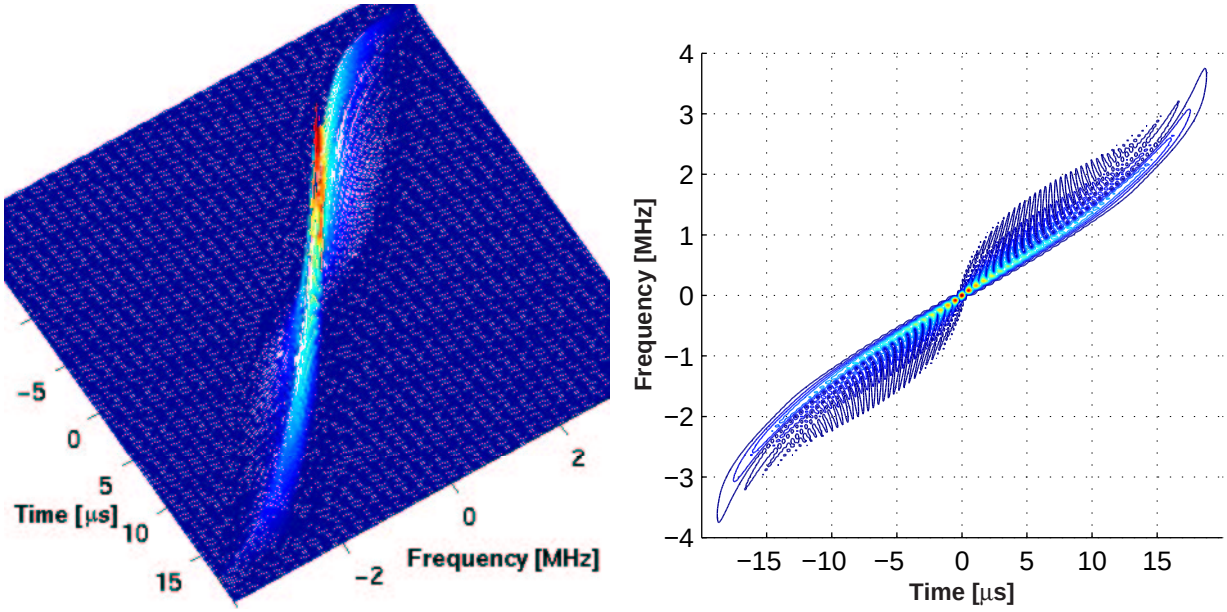

Figure 6.11: 3-D mesh and contour plot of the ambiguity function of the non-linear FM signal designed in 3.7.

6.4 Imaging with non-linear FM signals

The non-linear FM signal presented in Section 3.7 combine the attractive properties of high signal-to-noise ratio (17.5 dB) with relatively low sidelobes (-36 dB) without any weighting. The disadvantage is mainly the higher sensitivity of pulse compression of this signal compared to the linear FM in frequency shifts. This can be assessed by looking at the simulation results of Fig. 6.4, as well as by inspecting the ambiguity function of the non-linear FM signal shown in Fig. 6.11 and making a comparison with the ambiguity function of the linear FM signal shown in Fig. 2.5. In contrast to the linear FM signal, sidelobes appear everywhere away from the delay axis. There is a trade off in the non-linear FM signal design between compression robustness in frequency shifts and sidelobe levels. When the bandwidth of the amplitude spectrum, that the non-linear FM signal is matched with, is reduced, the resulting FM slope is closer to linear apart from the two ends at the beginning and end of the signal. This design results in a signal with similar properties as the tapered FM signal discussed in Chapter 4, where tapering is implemented using phase modulation rather than using amplitude modulation. Such a non-linear FM signal is less sensitive to attenuation, but has higher sidelobes similar to the unweighted FM case. Fig. 6.12 shows simulation results of imaging a number of point scatterers in the presence and absence of tissue attenuation. This design is a compromise between sensitivity in attenuation and sidelobe levels. Inspecting the echoes from the scatterer at depth 16 cm, for instance, it can be seen that in attenuating medium, the mainlobe widens and gets distorted around -22 dB, while the sidelobes increase by 5 dB, appearing now at about -31 dB.

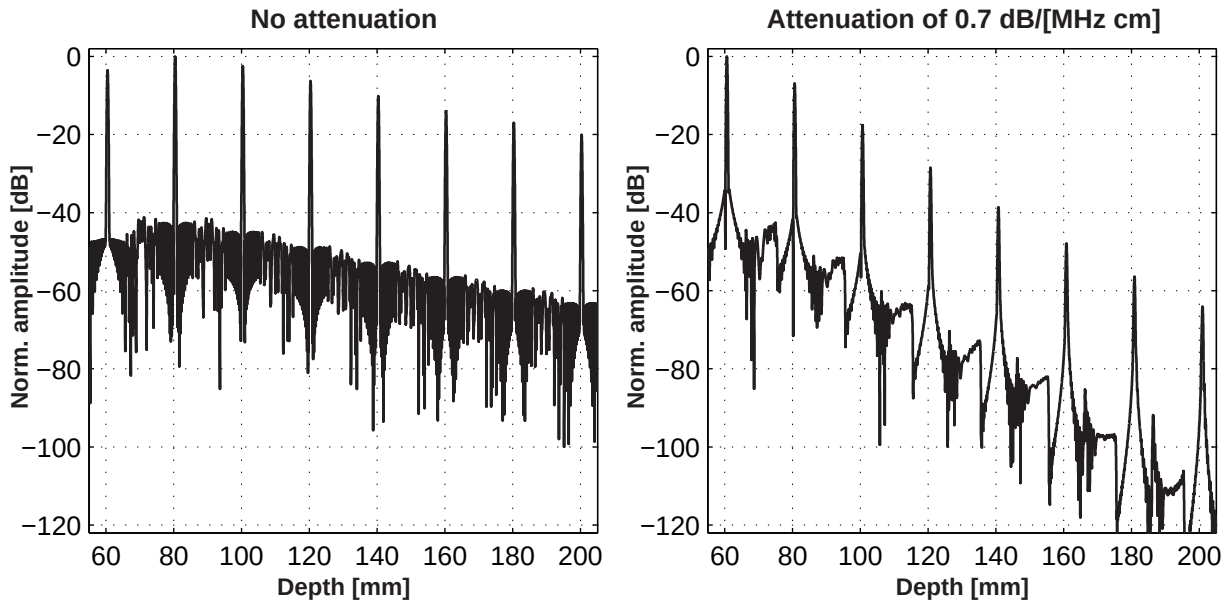

Figure 6.12: The effect of attenuation in sidelobe levels using non-linear FM excitation. Simulation results with Field II show the echoes from 8 point scatterers in the absence (left) and presence (right) of attenuation in the medium.

6.5 Imaging with complementary codes

Simulated images using Golay code excitation are shown in Fig. 6.13. Golay-coded imaging requires two transmit events for every line, which decreases the frame rate by half. Additionally motion artifacts are expected to degrade the sidelobe cancellation. The simulation results of Fig. 6.13 and Fig. 6.14 show that even in the case of imaging of stationary tissues, the complementarity property severely degrades in an attenuating medium. Perfect cancellation occurs when the medium has no attenuation Fig. 6.14a, while with an attenuation of 0.7 dB/[MHz$\times$cm], the range sidelobes increase up to -25 dB at a depth of 16 cm, in contrast to the tapered FM signal (Fig. 6.8), whose compression is very robust to attenuation. When the actual transducer impulse response is used in the simulations, the compression of Golay codes in the presence of attenuation is entirely unacceptable.

6.6 Evaluation of resolution and compression

Fig. 6.15 shows the response from a point scatterer for all coded waveforms considered in this chapter. The mainlobe width is a measure of the axial resolution. In the non-attenuating medium, the -20 dB mainlobe width from pulsed excitation is 1.49λ, for the tapered linear FM is 1.92λ, for the non-linear FM is 1.77λ, and for the Golay code is 2.17λ. That is, the mainlobe width of

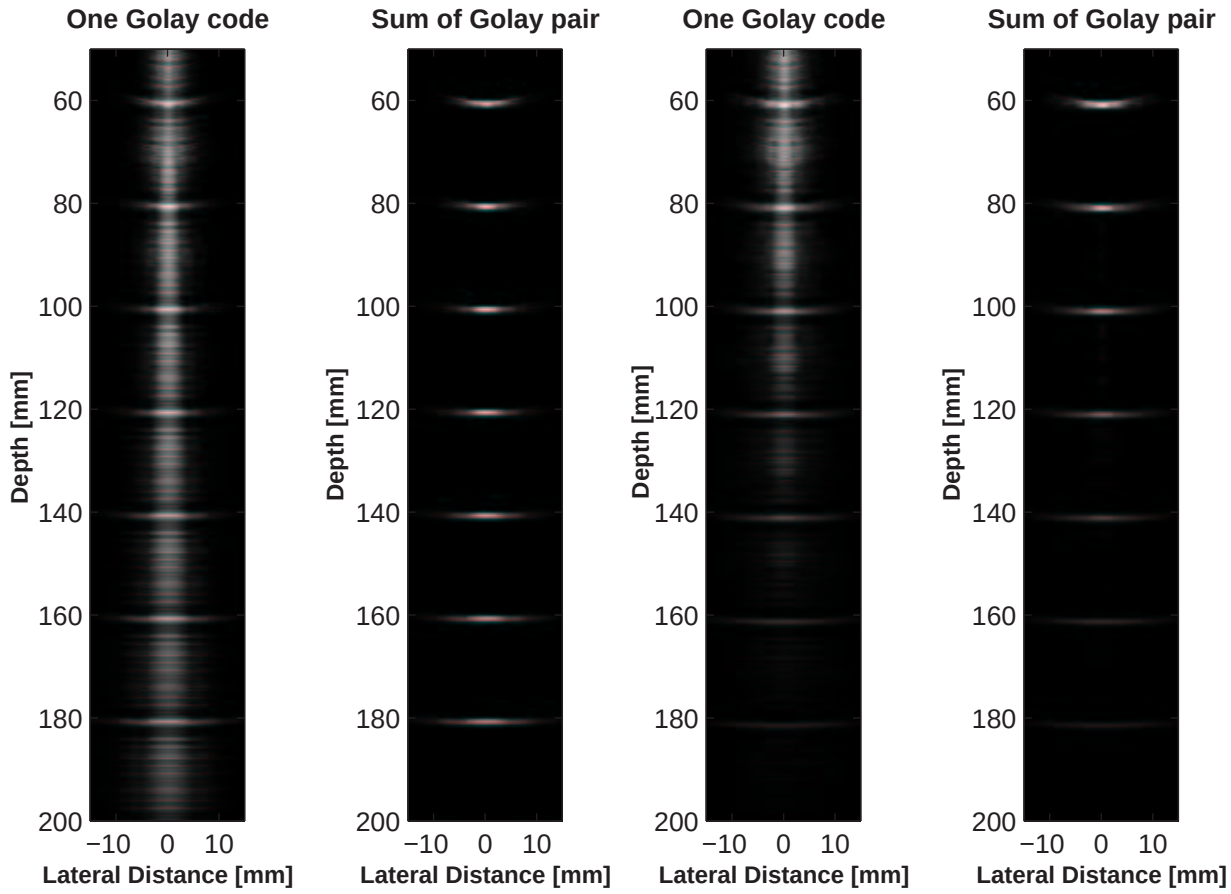

Figure 6.13: Imaging with complementary Golay codes in a medium with no attenuation (left two images) and with attenuation of 0.7 dB/[MHz×cm] (right two images).

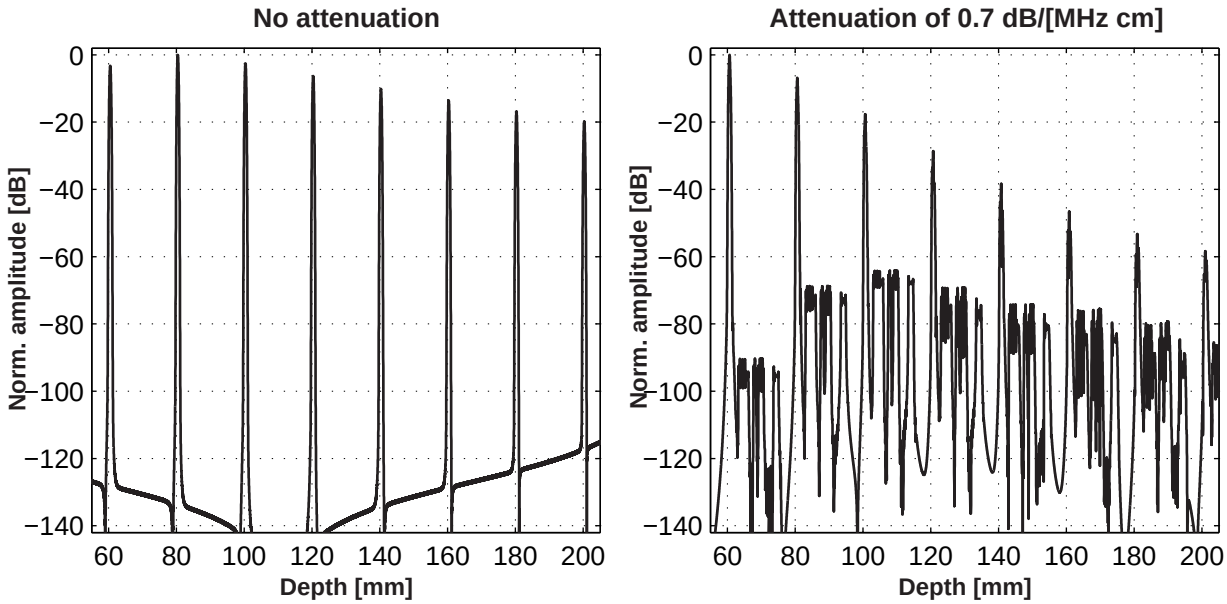

Figure 6.14: The effect of attenuation in sidelobe levels using complementary Golay codes. Simulation results with Field II show the echoes from 8 point scatterers in the absence (left) and presence of attenuation (right) in the medium.

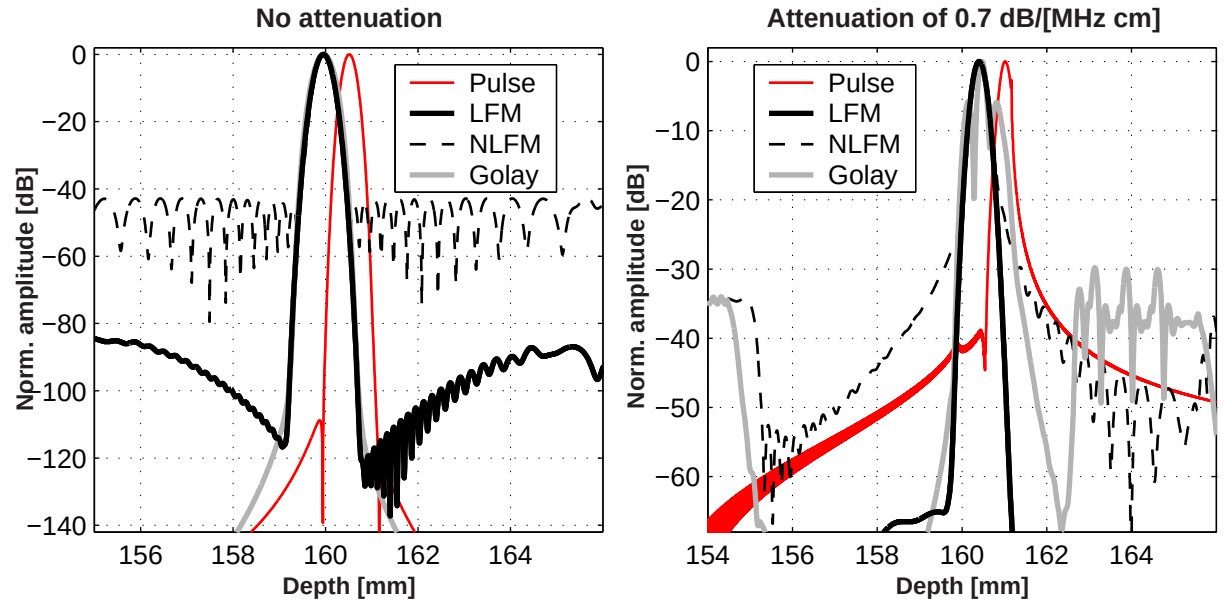

Figure 6.15: Response from a point scatterer positioned at depth of 160 mm for various coded excitation waveforms for the evaluation of axial resolution, in case of non-attenuation medium (left) and a medium with attenuation (right).

the tapered linear FM is 22% wider than that of the conventional pulsed excitation. The widening for the non-linear FM signal and Golay-coded excitation are 16% and 31% respectively. Imaging with the linear FM with the compression output shown in Fig. 4.6a results in axial resolution of 1.52λ, which is only 2% wider than that of the pulsed excitation. In this case, the sidelobes are slightly higher, especially in case of attenuation. The right graph of Fig. 6.15 shows the effect of attenuation on various excitation signals. The superiority of the tapered linear FM signal over all other signals is apparent. The -20 dB mainlobe width of the linear FM is 2.13λ, of the non-linear FM is 2.41λ, while the mainlobe of the Golay code is severely distorted and split into three lobes. The sidelobes for the linear FM signal are at -67 dB, while the sidelobes of the Golay codes are at only -30 dB. The energy outside the mainlobe for the linear FM signal is less than that of the pulsed excitation, i.e. image contrast will be improved using linear FM-coded excitation.

6.7 Pulse compression and array imaging

Coded excitation systems described in the literature use a single correlator at the output of the beamformer [10, 11]. This approach, although advantageous in terms of implementation, poses requirements on the code length and arises issues about the effect of time-gain compensation (TGC) and dynamic focusing on pulse compression. Dynamic focusing is a pixel-based non-linear operation, which will distort the phases of a long coded waveforms. In order to minimize this effect, O'Donnell [11] suggested that the code should have a duration less than $64/f_0$, which for the typical bandwidth of the ultrasound transducer restricts the time-bandwidth of the code to less than 30. Although this time-bandwidth product should give a SNR gain of 15 dB, the simulations presented in this chapter, showed that the transducer weighting and the attenuation significantly reduce the expected SNR gain. Additional mismatched filtering required for the high sidelobes of such small time-bandwidth product waveforms will shrink the SNR gain further. It is, therefore, questionable if such a coded excitation system can have any practical benefit.

In a fully-digital system, **pulse compression should precede and beam formation should follow**. It has been shown that the compressed rf-data resulting from coded excitation do not differ significantly from conventional rf-data using pulse excitation. In this way, there are no requirements for the code length in terms of focusing and time-gain compensation. Long codes can be used, which will result in good compression and SNR gain. The disadvantage of this, is that compression has to be done in all transducer channels, which requires one compression filter on every channel.

Coded excitation could also be used for flow estimation. Flow estimation is one of the areas that could actually benefit the most from coded excitation. This is because the signals from blood are usually very weak; sometimes more than 40 dB below the surrounding tissues. The higher signal-to-noise ratio achieved with coded excitation will improve the estimator's accuracy. Applying a flow estimator before compression could also be possible, but the effect of the varying frequency of the signal on the estimator has to be investigated. Thus, it is preferable to compress first if

implementation allows, and apply one of the traditional flow estimators afterwards.

CHAPTER

SEVEN

Clinical evaluation of coded imaging

In this chapter acquired images using coded excitation are presented.

7.1 Experimental setup

The experimental system used was based on a commmercial scanner (B-K Medical 3535) shown in Fig 7.1 with a single-element mechanically-rotating transducer. The transducer is motor-driven and scans a sector as shown in Fig 7.2. The scan angle is fixed at 75 degrees and the pulse repetition frequency can be set in six modes from 1 to 3.4 KHz. Depending on the settings for imaging depth, the machine can acquire from 34 to 212 number of lines. The machine was set for most experiments to a mode for acquisition of 106 lines, scanning a depth from 0 to 180 mm with a pulse repetition frequency of 3 KHz.

The transducer crystal has a nominal frequency of 4 MHz. The impulse response of the transducer was measured and it is shown in Fig 7.3. The -6 dB bandwidth was 65%.

The single-element module on the scanner was modified in order to operate with external excitation. The transmitter was inactivated and was interfaced to an external pulse generator. During the research project, two different experimental set-ups were used.

In the first system, a LeCroy 9112 arbitrary function generator (AFG) was used for the generation of coded waveforms on the computer. A software library was written for the control of the AFG through MATLAB^{TM} [57]. A power RF amplifier (RITEC 5000), specifically designed to drive ultrasound transducers, was used subsequently for amplification. The acquisition of the RF data was done using the sequential acquisition capabilities of an 8-bit digital oscilloscope (LeCroy LC334AM) at a minimum sampling frequency of 20 MHz depending on the requested depth of penetration and number of scan lines. The transfer time of raw binary data for an image with 106 scan lines and a depth of 18 cm was 0.22 s, allowing acquisition rate of 4.5 frames/sec. The total processing and display time was 2.714 s/image. A block diagram of the system is shown in Fig 7.4.

Figure 7.1: The ultrasound scanner (B-K Medical Model 3535) used in the experiments.

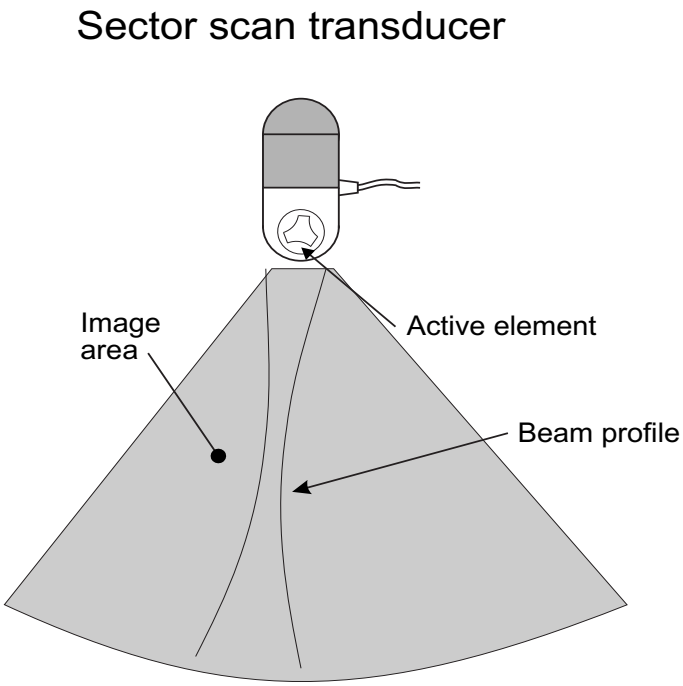

Figure 7.2: The single-element transducer (B-K Medical) which makes sector images by mechanical rotation. The sketch is taken from Jensen [2] after permission.

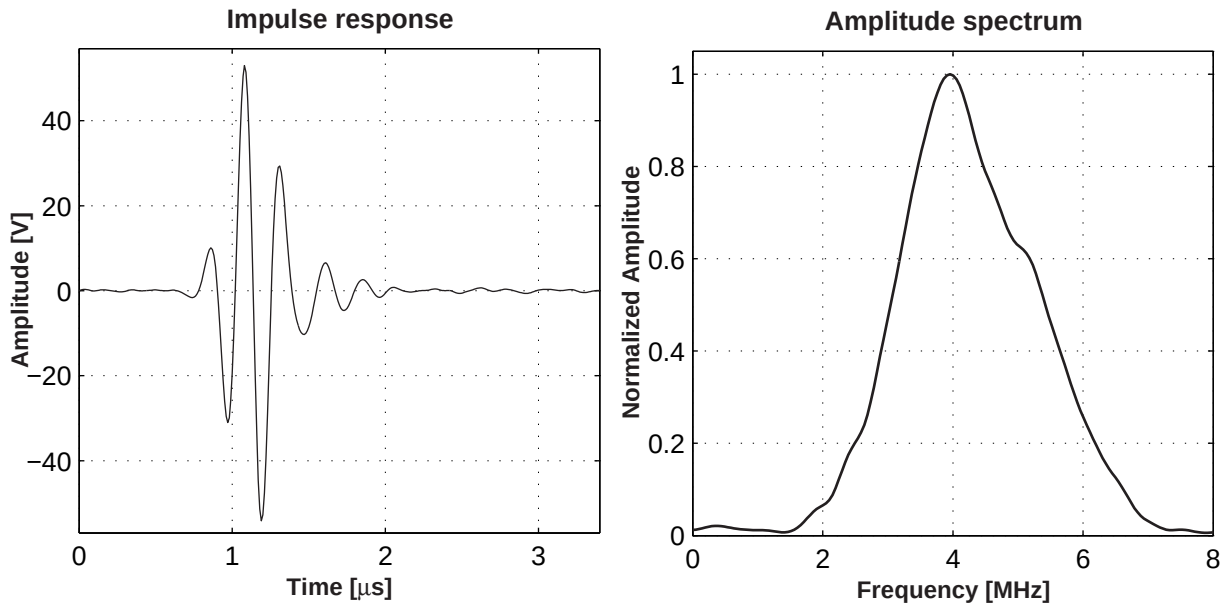

Figure 7.3: Measured impulse response and transfer function of the single-element transducer (B-K Medical) used in the experiments.

Both the generation of the signals through the AFG and the image acquisition with the oscilloscope were controlled remotely from the computer through a GPIB interface. A software library for the communications with the oscilloscope was written, allowing high level programming through MATLAB^{TM}. After acquisition of the RF data, all post processing and display is done on the computer. That includes high-pass filtering, pulse compression, interpolation and scan conversion, logarithmic compression, envelope detection and display. The whole system was running under Linux and could be operated from any computer connected to the network.

In the second system, the same scanner was interfaced to one transmitter and one receiver board from the Center's newly constructed RASMUS experimental system [3]. A picture of this system is shown in Fig 7.5. The software control of both boards is implemented as a MATLAB^{TM} toolbox of high-level commands written by Svetoslav Nikolov. The transmitter board is capable of transmitting different complex arbitrary waveforms for each line of an image with a few lines of MATLAB^{TM} code. The board was programmed to allow alternating excitation on every second frame. That allowed direct comparison of the same set of image pairs; one with conventional and one with encoded excitation, or image pairs with complementary sequences. The analog RF data of the scanner were sampled at 40 MHz. The receiver sampling unit has 12-bit ADCs and 2 GBytes storage SDRAM. That allowed storage of 140 consecutive ultrasound images, corresponding to about 25 s of scan data. Subsequently the stored RF data were read from the memory, and all post processing (pulse compression, interpolation, scan conversion and display) was done off-line on the computer as in the first system.

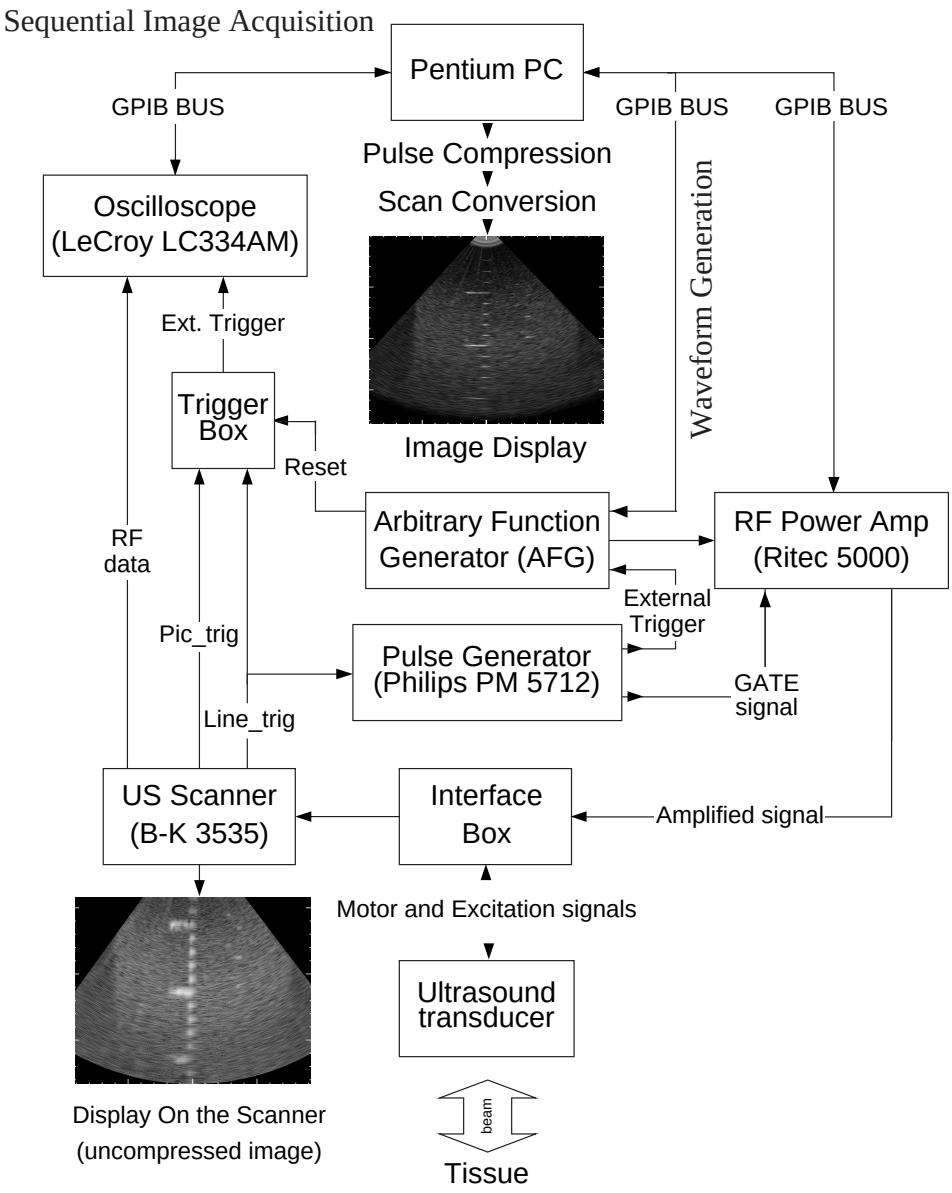

Figure 7.4: The first experimental system based on an arbitrary function generator and on a digital oscilloscope for image sampling.

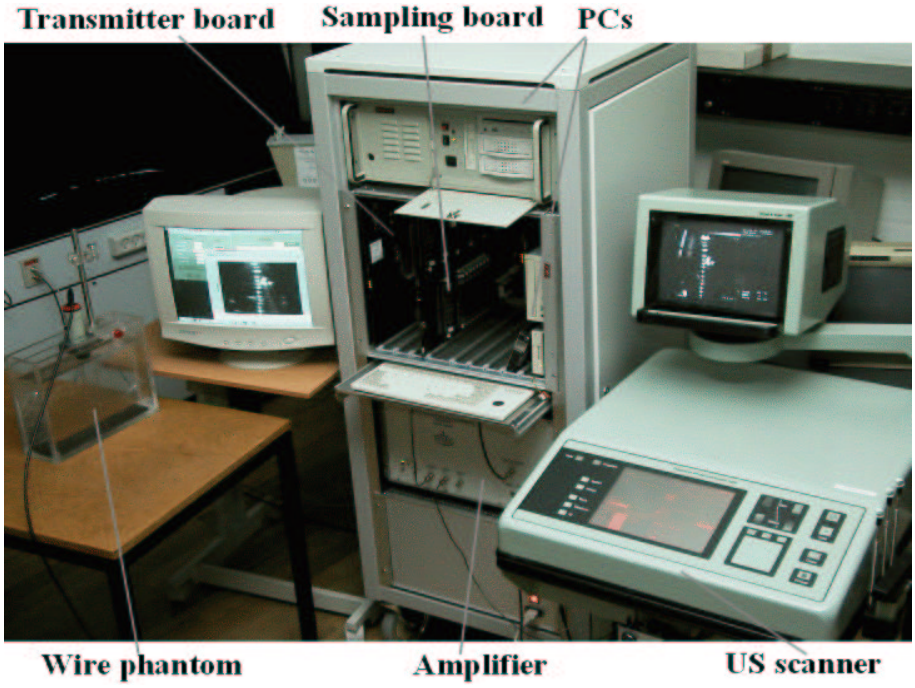

Figure 7.5: The second experimental system using boards from the Center's newly constructed RASMUS experimental system [3].

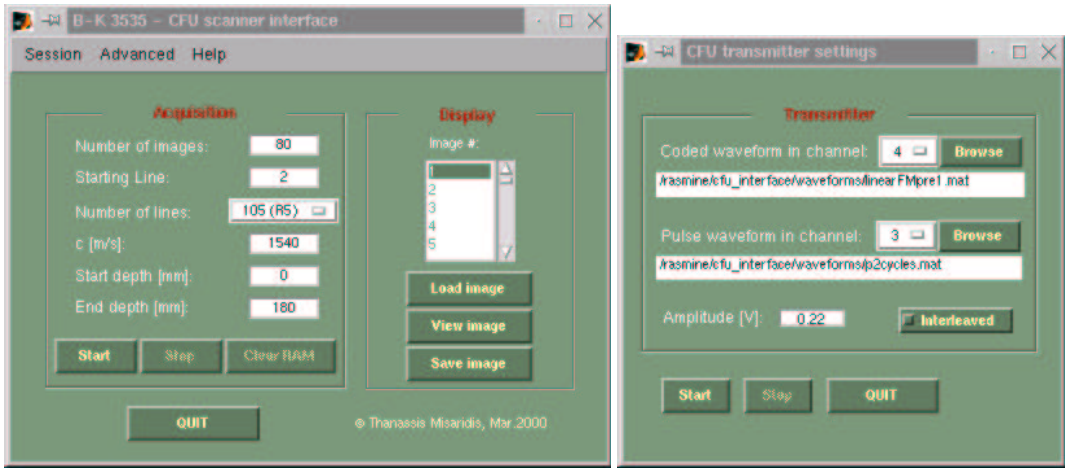

Figure 7.6: Graphical user interface in MATLAB^{TM} for imaging setup and acquisition with coded excitation.

7.2 Phantom images with coded excitation

Initially, pulse-echo measurements from a wire phantom in water were performed. It has not been possible to evaluate the sidelobe level of the compressed output for the proposed code, but we can conclude that the sidelobes are below -50 dB, which was the acoustic noise for the pulsed excitation output. The acoustic noise was due to reverberation at the water surface and the bottom of the tank. The -6 dB axial resolutions were measured and were found to be 0.94λ for the coded and 1.18λ for a two-cycle pulsed excitation.

In this section images of wire phantoms are presented. The first set of images shown in Fig. 7.7 have been acquired with the first experimental system, where sampling was done by the 8-bit digital oscilloscope. The available 8 bits of the sampled returned echoes limit the dynamic range to 48 dB, while sampling noise is considerable at high depths. Fig. 7.7 shows images of a phantom with attenuation of 1 dB/[MHz×cm]. The peak excitation voltages are 32 V p-p for the conventional pulse and 20 V for the chirp, that yield the same I_{sptp}. The applied excitations are very low compared to what is commonly used in ultrasound scanners. Therefore, the noise level in the conventional image of Fig. 7.7 is high. However, when a chirp excitation is applied, there is a significant reduction in noise level for depths over 10 cm. Some RF lines were corrupted with spike noise resulting in brighter stripes along the image. It was found that there was a sampling error in the oscilloscope at a certain sampling frequency and this frequency was not used subsequently. The spike noise was removed by a median filter whenever it occurred. The time-gain compensation was kept low and at the same level for both images. The -20 dB axial resolution, measured at the scatterer located at about 6 cm, was 0.71 mm for the conventional pulse excitation and 1.04 mm for the chirp. In the latter, the axial resolution depends on the applied compression filter and can be improved at the expense of higher axial sidelobes.The dynamic range of the images shown in Fig 7.7 is 50 dB.

A second set of phantom images have been acquired using the second experimental system and are shown in Fig. 7.8. The higher dynamic range of the 12-bit sampling board of the RASMUS system [3] resulted in better quality images. A different phantom was used with attenuation of 0.5 dB/[MHz×cm]. This phantom consisted of wires of 0.2 mm in in diameter, positioned every 1 cm axially. Additional wires were placed at a 15 degree angle with decreasing distance down to 1 mm.

In the pulsed image of Fig. 7.8 a two-cycle pulse of the carrier frequency was used. In the coded image the excitation was a 20 μs tapered FM signal with the theoretical compression output shown in Fig. 4.6. The -6 dB axial resolution of the coded image measured at the wires in depth 12 and 14 cm was 1.31λ. The resolution of the pulsed image was 1.49λ. In conventional imaging, short broadband pulses have to be used in order to achieve good axial resolution and to use all available system bandwidth, with the drawback of degrading the SNR. When FM signals are transmitted, the axial resolution depend on the bandwidth that is swept and not on the duration of the signal. Therefore, coded images using FM modulated signals result not only in a better SNR, but also in

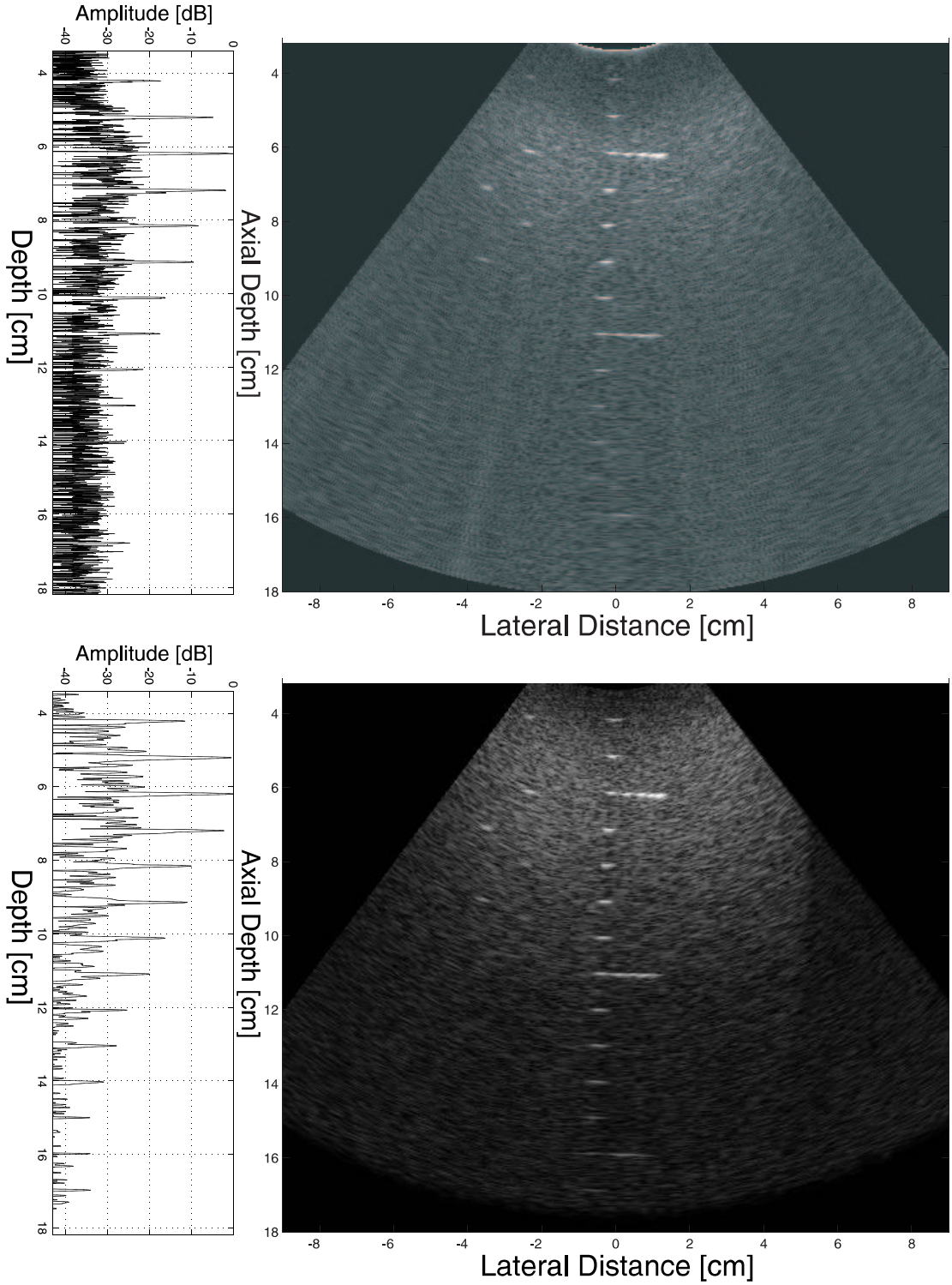

Figure 7.7: Images of a wire phantom with attenuation of 1 dB/[MHz×cm]. The dynamic range of both images is 50 dB. The peak excitation voltages 32 V for the conventional pulse and 20 V for the chirp. The plots on the left side are the central RF lines of the images.

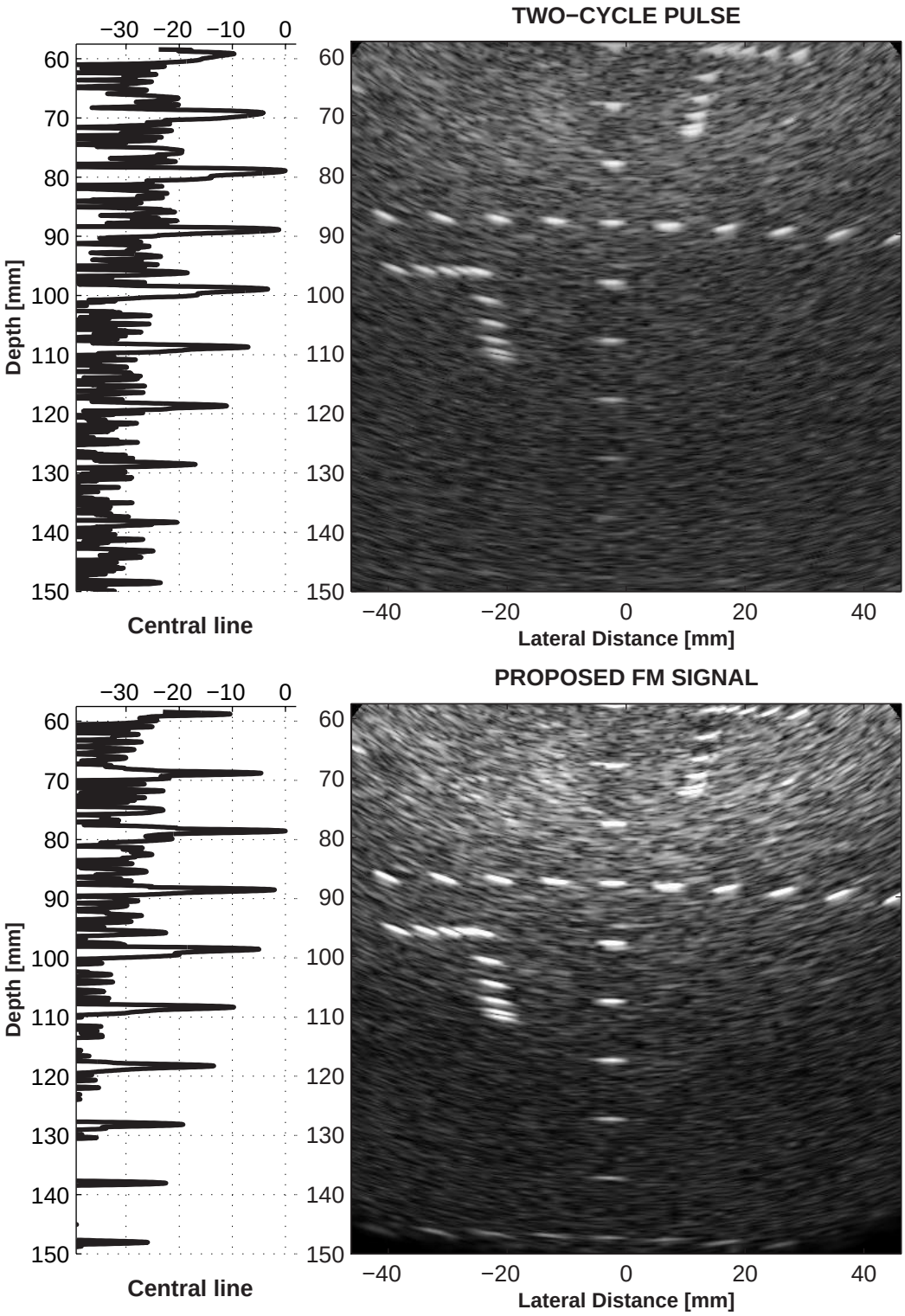

Figure 7.8: Detail of images of a wire phantom (right) and central RF lines (left) for coded and pulsed excitation. Matched filtering has been applied to both images. The dynamic range of the images is 45 dB. From the graphs on the left, an improvement in SNR of about 10 dB can be seen. Axial resolution is also higher for the coded image.

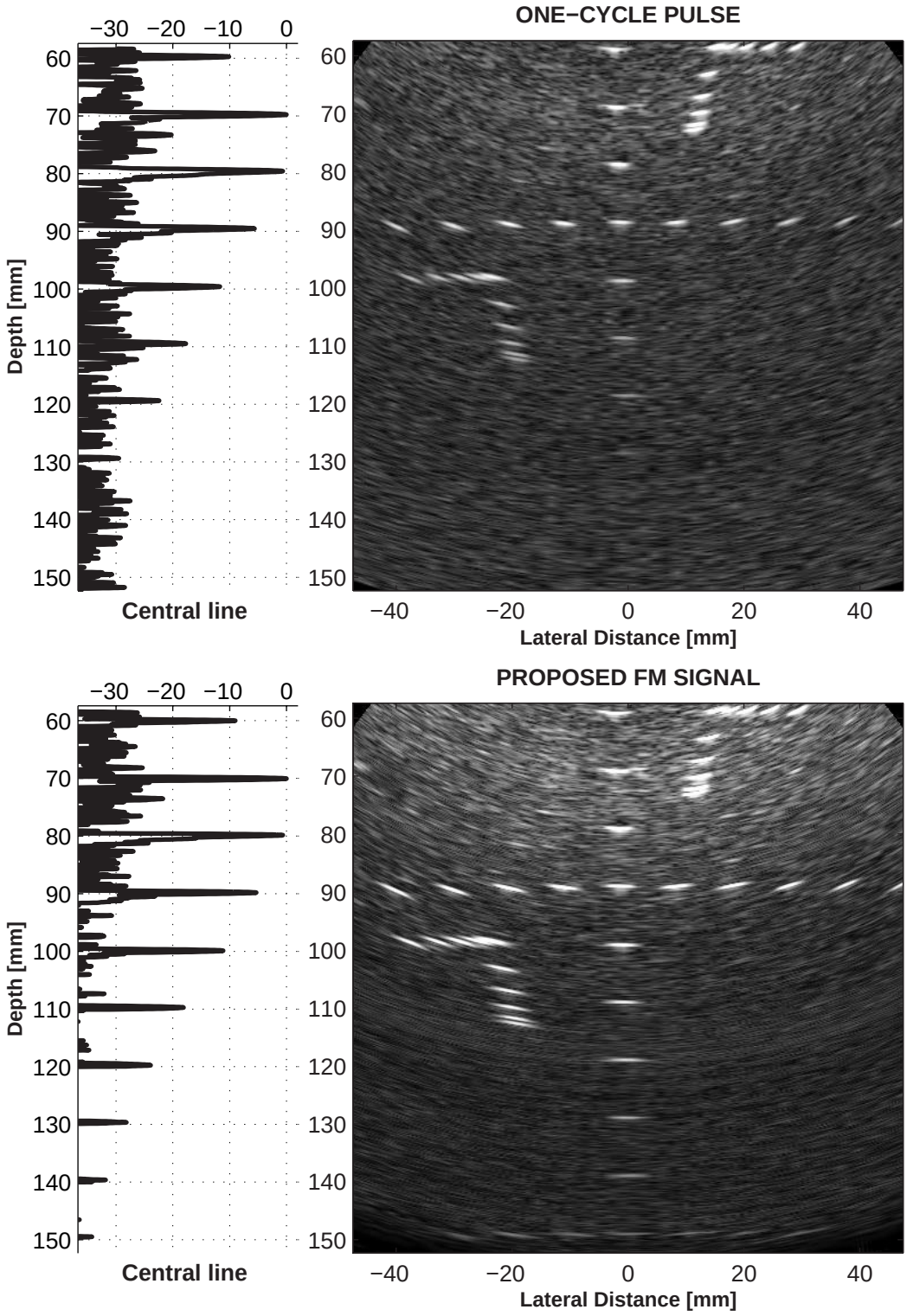

Figure 7.9: Another set of images of a wire phantom (right) and central RF lines (left) for coded and pulsed excitation. Matched filtering has been applied to both images. The dynamic range of the images is 45 dB.

better axial resolution by utilizing the available bandwidth more efficiently. Fig. 7.8 shows that there is an effective gain in SNR of 10 dB or more, that corresponds to an additional penetration of 3-4 cm with the 4 MHz probe.

Another set of images from a wire phantom is shown in Fig. 7.9. In the pulsed image of Fig. 7.8 a one-cycle pulse of the carrier frequency was used. In the coded image the excitation was a 20 μs tapered FM signal with the theoretical compression output shown in Fig. 4.6b. Images of the same phantom using a complementary pair of Golay codes with length 40 are shown in Fig. 7.10. A single Golay code has high axial sidelobes that are visible on the left image. When the echoes from 2 Golay codes are added coherently, there is a degree of cancellation on the axial sidelobes. However, confirming the simulation results, the cancellation of the sidelobes is not perfect due to attenuation even for the stationary phantom imaging, and shadows are still visible along the wires.

7.3 Clinical images with coded excitation

Clinical images of the abdomen using the proposed scheme are shown in Fig. 7.11 and Fig. 7.12. The images show an excellent performance of the encoded excitation in terms of noise reduction at large depths and resolution. The auto-covariance matrix on the image gives an indication of the speckle size. The lateral resolution of speckle data taken from the images of Fig. 7.12 is very similar for both images, while axially the coded image has slightly better performance (Fig. 7.13).

These results clearly demonstrate that abdominal ultrasound imaging can benefit from coded excitation yielding a higher SNR and therefore penetration, while maintaining both axial and lateral resolution. The higher SNR can be exchanged with resolution by increasing the center frequency, i.e. for GSNR=10 dB, going from 4 to 5 MHz without compromising SNR. Longer codes can make this frequency step even bigger.

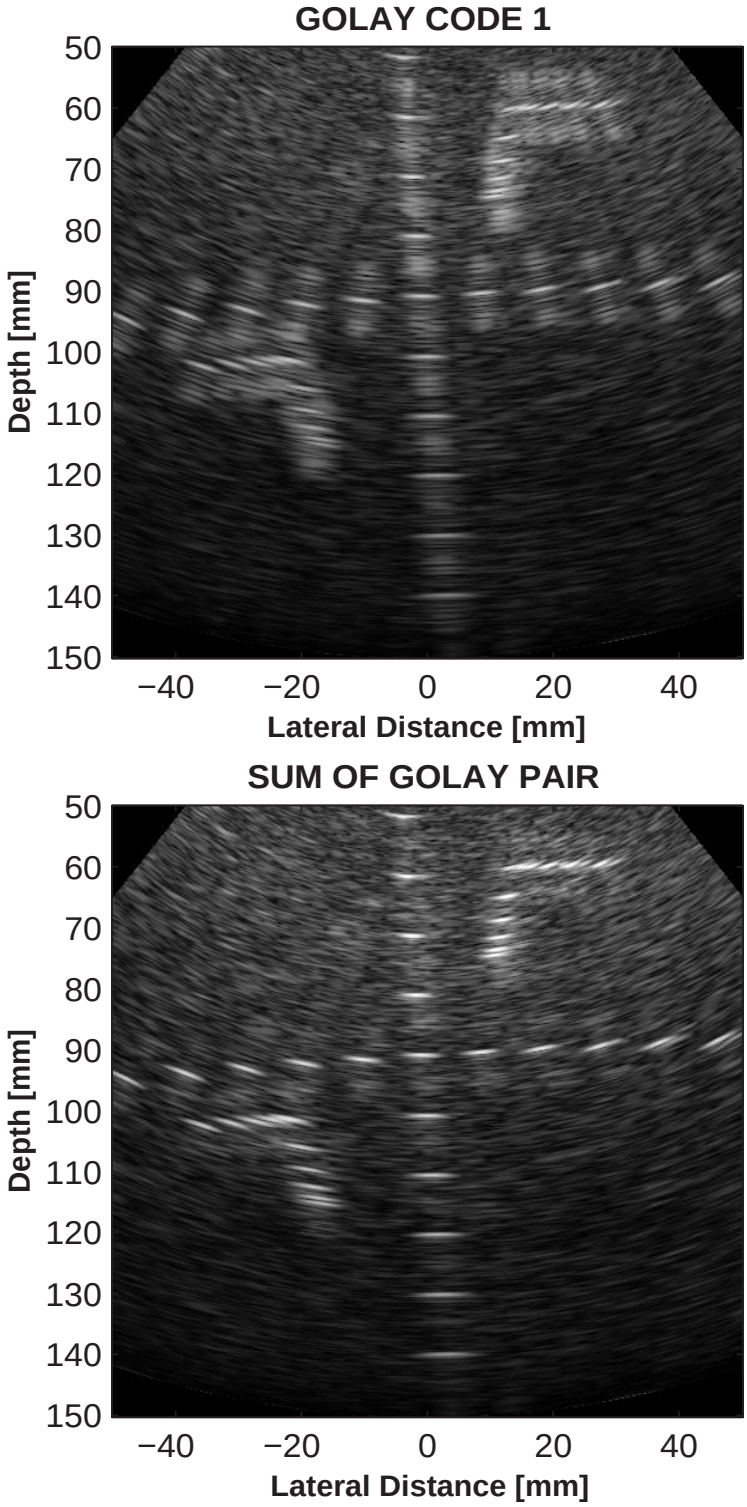

Figure 7.10: Images with Golay pair excitation of a wire phantom with attenuation of 0.5 dB/[MHz×cm]. On the left is the image with one of the Golay codes and on the right is the sum of the two complementary images. The dynamic range is 45 dB.

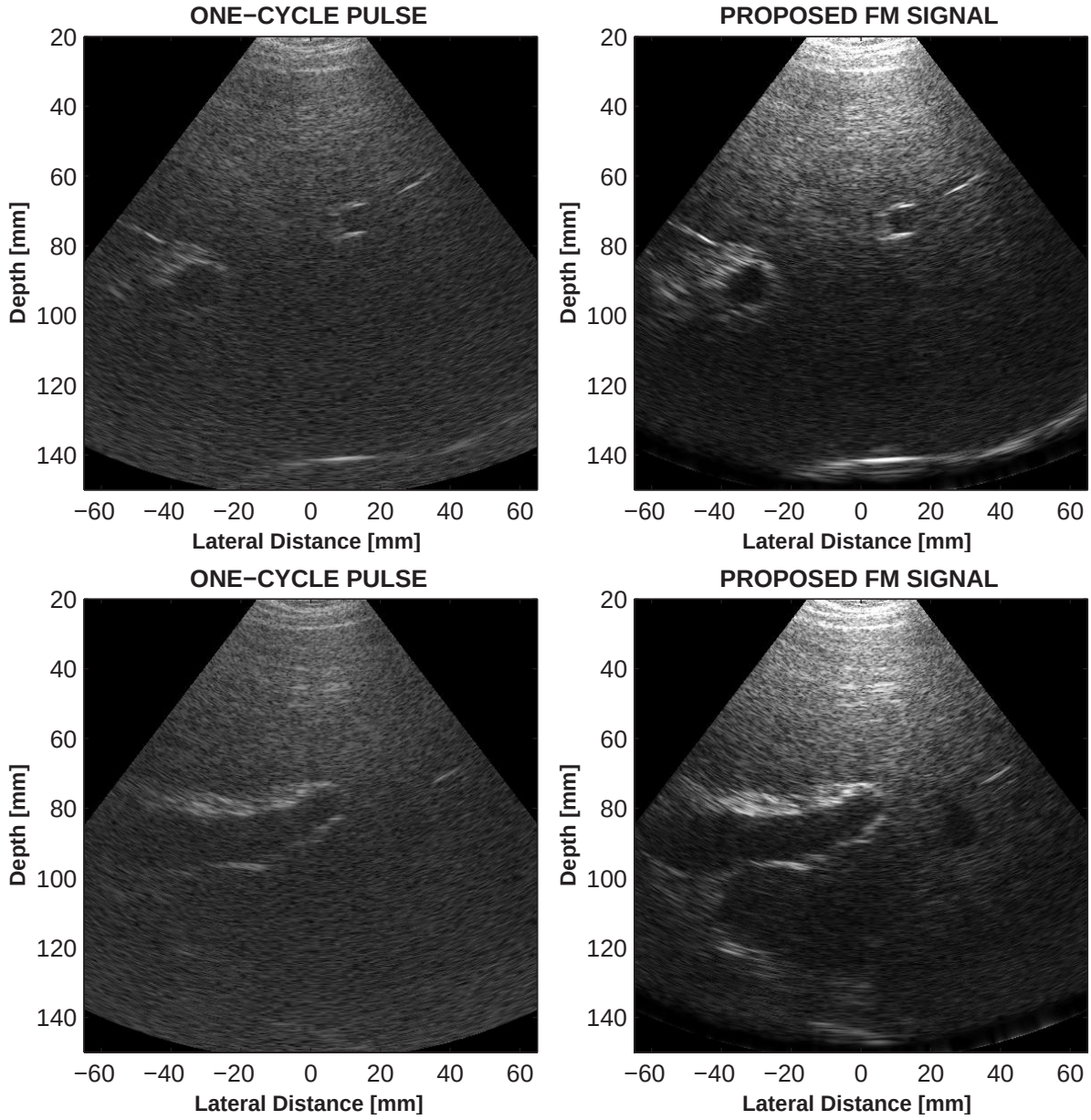

Figure 7.11: Clinical images with linear FM excitation On the left is the image with one of the Golay codes and on the right is the sum of the two complementary images. The dynamic range is 45 dB.

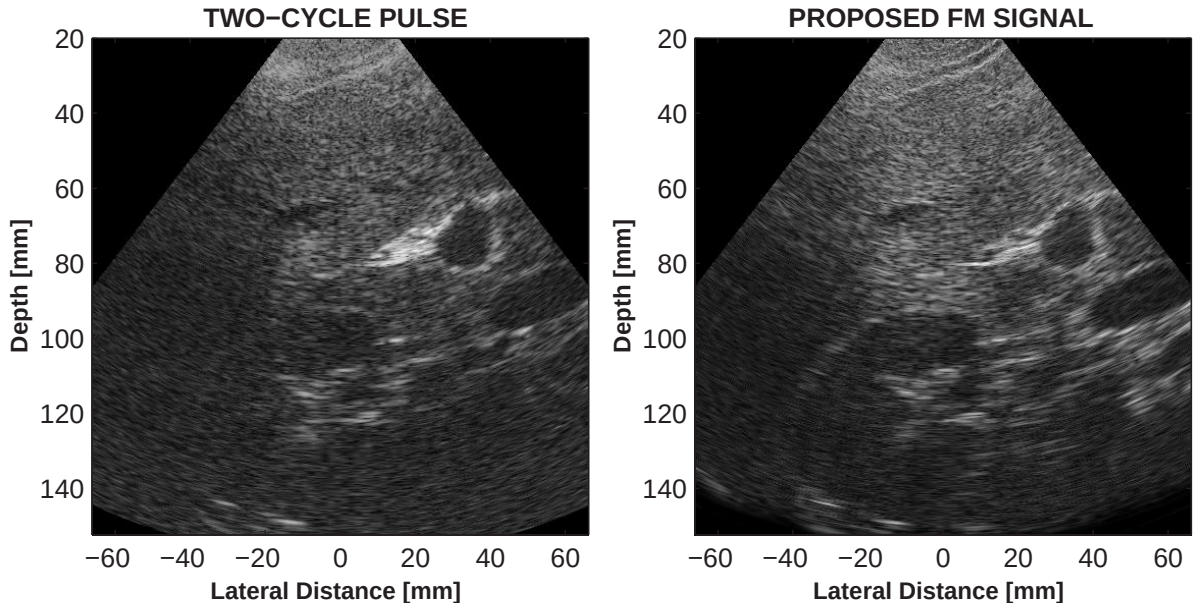

Figure 7.12: Clinical images of the right kidney for coded and pulsed excitation. The portal vein and the inferior vena cava are at the right side of the images and liver tissue is left from the kidney. The dynamic range of the images is 45 dB. Improvement in resolution and noise reduction at large depths are visible.

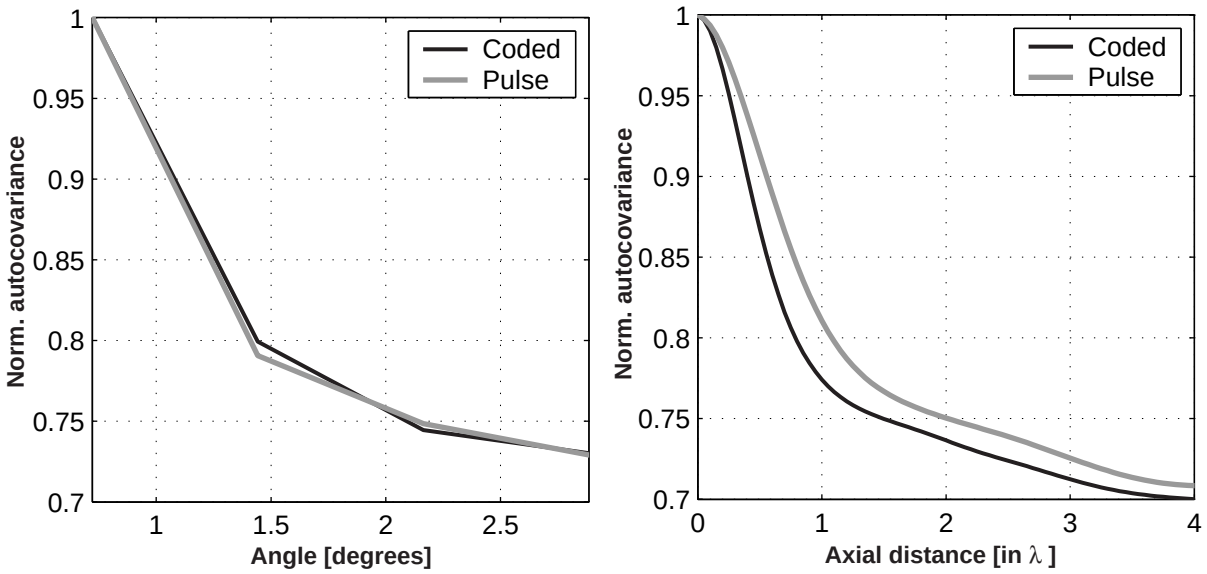

Figure 7.13: Evaulation of the lateral (left) and axial (right) resolution in speckle using autocovariance matrix analysis. Speckle data are taken from the images of Fig. 7.12. The gray lines correspond to the pulsed image.

CHAPTER

EIGHT

Waveform diversity for fast ultrasound imaging

The second objective of coded excitation (with the improvement in signal-to-noise ratio in ultrasound images being the first) is the possibility of simultaneous transmission of a number of coded signals, which can be distinguished on the receiver. Coded excitation has, thus, the potential of increasing the frame rate in ultrasound imaging using parallel transmission. For good axial resolution, all signals in the set should have good auto-correlation properties. Since a correlation receiver is used for pulse compression, separation of the echoes from all excitation signals requires that the cross-correlation of any two signals in the set is low for all relative phase shifts. The cross-talk in the receiver results when the filter matched to one of the signals is applied to the echoes originating from another transmitted signal in the set.

The arising issues, which are the topic of the second part of the dissertation, are i) the orthogonality among the transmitted signals that can be achieved and ii) how these signals can be used, i.e. coding imaging strategies.

8.1 Waveform diversity for the FM signal

In Chapters 3 and 4 the properties of the linear FM signal and its matched and mismatched response have been analyzed. This corresponds to the auto-correlation and a weighted auto-correlation function of this important waveform. What is of interest for fast imaging, is the FM signal diversity, i.e. the number of FM signals that can be constructed and used for ultrasound imaging, as well as the cross-correlation properties of these signals.

The diversity of the FM signal is far more limited than that of the digital phase codes. The design parameter of the FM signal is the FM slope μ, which is associated with the duration T and the bandwidth B. Fig. 8.1 illustrates two linear FM signals having different slopes μ_n and μ_m. Cook

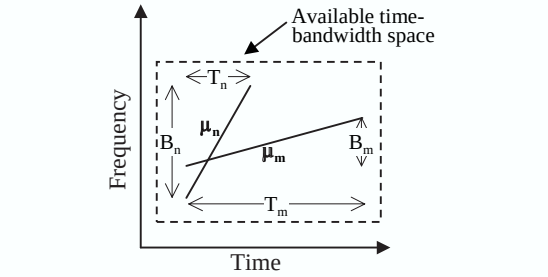

Figure 8.1: Diagram showing two linear FM signals with different FM slopes $\mu_n = B_n/T_n$ and $\mu_m = B_m/T_m$ and the same time-bandwidth product $T_nB_n = T_mB_m$.

and Bernfeld [25] have calculated analytically the mismatched filter response, which occurs when the transmitted signal has an FM slope of μ_n and the slope at the receiver is μ_m, i.e. the cross-correlation between FM signals with different slopes. Defining the *mismatch factor* γ as:

$$\gamma = \frac{\mu_n - \mu_m}{\mu_n}, \tag{8.1}$$

the magnitude of the mismatched output is [25]:

$$|g(t)| = \frac{1}{\sqrt{2\,|\gamma|}} \left\{ [C(Y_1) + C(Y_2)]^2 + [S(Y_1) + S(Y_2)]^2 \right\}^{1/2}, \tag{8.2}$$

where the arguments Y_1 and Y_2 of the Fresnel integrals are:

$$Y_1 = \sqrt{\frac{\gamma\mu_m}{2}}\, T_m \left(1 - \frac{2t}{\gamma T_m}\right) \quad , \quad Y_2 = \sqrt{\frac{\gamma\mu_m}{2}}\, T_m \left(1 + \frac{2t}{\gamma T_m}\right) \tag{8.3}$$

Due to the Fresnel structure of (8.2), the cross-correlation function of FM signals with different FM slopes resemble linear FM spectra of low time-bandwidth product. The Fresnel term $\left\{ [C(Y_1) + C(Y_2)]^2 + [S(Y_1) + S(Y_2)]^2 \right\}^{1/2}$ is approximately equal to $\sqrt{2}$ for $t = 0$. Then the ratio of the amplitude of the distorted mismatched filter output to that of the ideal matched filter is [58]:

$$\frac{(Mismatched \quad filter \quad output)}{(Matched \quad filter \quad output)_{\max}} = \frac{1/\sqrt{\gamma}}{T_m\sqrt{\mu_m}} = \frac{1}{\sqrt{\gamma T_m B_m}}. \tag{8.4}$$

Fig. 8.2 shows the auto- and cross-correlation functions for two tapered linear FM signals, which have the same time-bandwidth product of 67 and a mismatch factor γ=0.84. From (8.4), the cross-talk between the signals will be $20\log(\sqrt{0.84 \cdot 67})$=17.5 dB below the auto-correlation peak.

Although such cross-talk level is sufficiently low for good separation in ultrasound imaging, as it will be shown in the following chapter, the auto-correlation functions have high sidelobe levels, which are inappropriate for imaging. Unfortunately, when the usual approach of weighting is

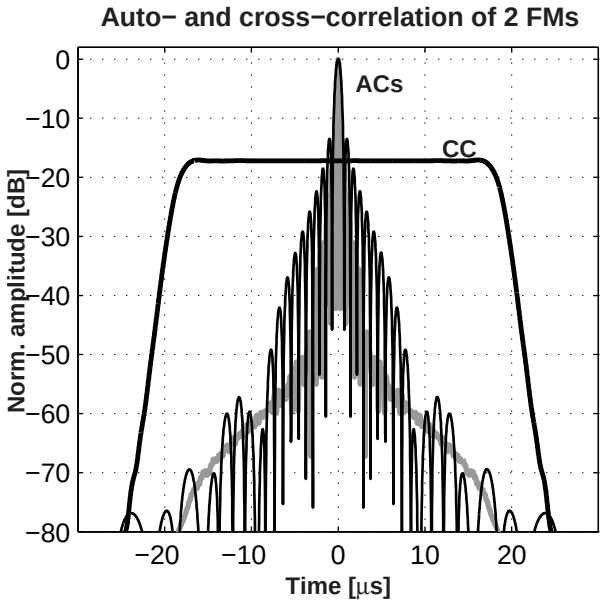

Figure 8.2: Auto- and cross-correlation functions for two tapered linear FM signals, one with T=10 μs and B=6.7 MHz and the other with T=25 μs and B=2.7 MHz. The two signals have the same time-bandwidth product of 67 and a mismatch factor γ=0.84.

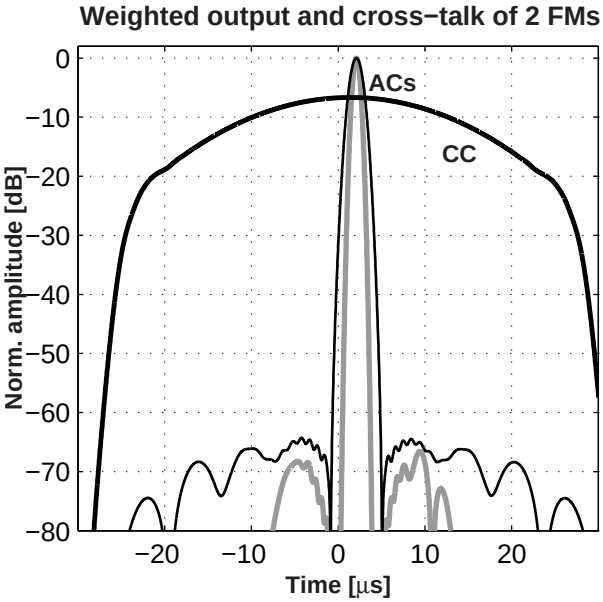

Figure 8.3: Compression output and cross-talk for the two tapered linear FM signals, when weighting is applied on the receiver filters for sidelobe reduction.

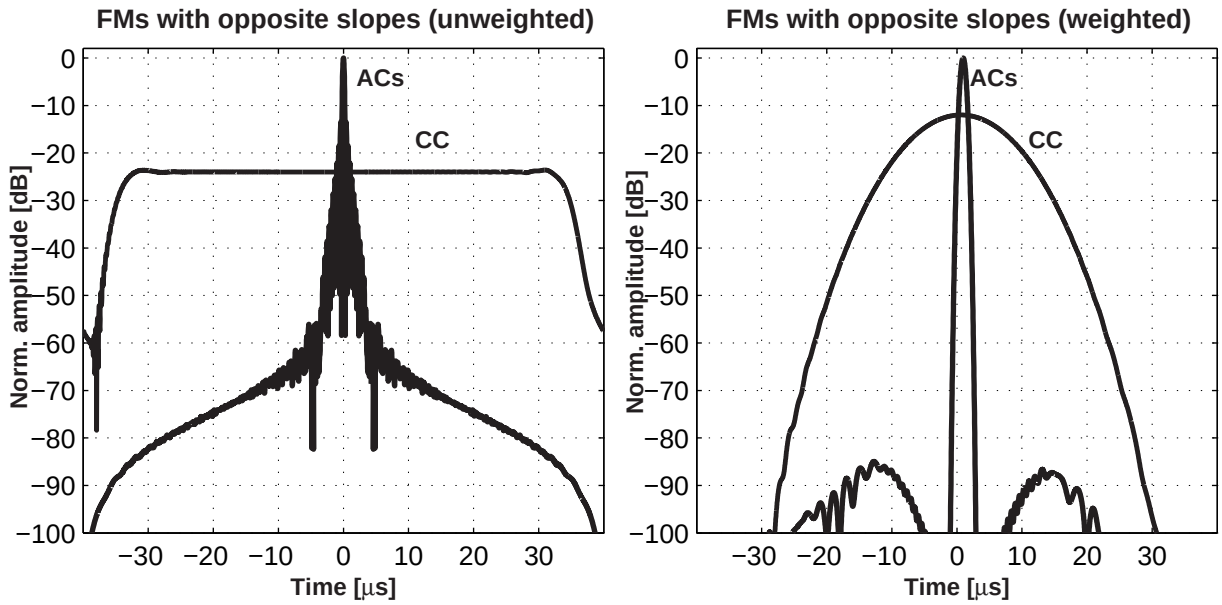

Figure 8.4: Compression output and cross-talk for the two tapered linear FM signals with equal and opposite FM slopes. The first design has minimum cross-talk and the second has minimum axial sidelobes.

applied to the receiver filters, the cross-talk increases from -17.5 dB to -7 dB (Fig. 8.3). For limited energy signals, the area under the squared magnitude of both the auto- and the cross-correlation functions is constant. Therefore, it can be reasoned [59], that the shape of the cross-correlation which gives the lowest values should be a constant function. Mismatching of the receiver does not relax the volume constrains of the auto- and cross-correlation functions [1], and any weighting will yield higher cross-correlation than the constant unweighted cross-correlation shown in Fig. 8.2. Fig. 8.3 shows the results when the input signals are cross-correlated with weighted compression filters. In this case, the axial sidelobes have diminished, but the cross-talk has increased by 10 dB. In conclusion, the cross-correlation among FM signals with different slopes is low, if no weighting is applied. A set of such FM signals can be used in communication systems as carriers for binary data [58], where axial sidelobes are not a problem. Using such set for imaging requires weighting, which in turn increases the cross-talk among the signals in the set.

Another drawback of this design is that the two or more signals have different duration and bandwidth. The duration is associated with the transmitted energy and the bandwidth is associated with the axial resolution of the image. When such set of FM signals are transmitted simultaneously from different elements of an ultrasound array transducer, resolution and heating effects will not be uniform across the image.

The obvious choice, when only two linear FM signals are required, is two signals with slopes equal in magnitude and opposite in sign. Such signals have the same duration, and sweep the same

bandwidth one with increasing and the other with decreasing frequencies. Both signals can utilize the entire available time-bandwidth space, which result in the best resolution and highest SNR gain. This design also yields the highest possible value of 2 for the mismatch factor γ, and, thus, the highest orthogonality among two FM signals.

Fig. 8.4 shows the compression output and cross-talk for such a pair of FM signals, 20 μs in length, with 6.9 MHz bandwidth and a time-bandwidth product of 140. When no weighting is applied at the receiver (Fig. 8.4a), the cross-talk (cross-correlation function) is constant with amplitude of $20\log(\sqrt{2\cdot 140})$=24.4 dB below the auto-correlation peak. For weighted receiver filters (Fig. 8.4a), the cross-talk is -12 dB. Fig. 8.4 shows the trade-off between cross-talk and axial sidelobes in the useful signal. For the given time-bandwidth product of 140, the cross-talk between two FM signals lies between -12 dB and -24.4 dB, while the axial sidelobes lie between -13 dB and -85 dB. Different weighting on the receiver can yield any values in this range, according to the specific imaging needs.

8.2 Frequency division

Another possibility is to divide the available bandwidth in two frequency bands and let each FM signal sweep its own band. Since the actual transmitted signals are their convolution with the transducer impulse response, both signals should have sufficient energy close to the center frequency of the transducer. In this way, the SNR gain can be kept high. On the other hand, cross-talk is reduced only if the signal spectra do not overlap. These contradictory requirements limit the method. In Fig. 8.5 the design of such signals and their correlation properties are shown. They are based on simulation of a transducer at 8 MHz with a 65% fractional bandwidth and two broadband preweighted FM signals centered at 6 and 10 MHz. The spectra of the two signals are shown at the left and the auto- and cross-correlation are shown at the right of Fig. 8.5. The different FM slopes allow separation of about 12 dB between the two signals. Frequency division is responsible for further cross-talk reduction. The axial sidelobes are less than -70 dB, and the cross-talk between the signals is around -60 dB. The axial resolution is inversely proportional to the bandwidth of the coded signals, and therefore with this method the axial resolution degrades by about 80% compared to the axial resolution of one signal occupying the entire available bandwidth. A loss in SNR is expected since the transmitted energy is not centered at the transducer central frequency.

8.3 Cross-correlation (CC) of binary codes

When only two sequences are needed, m-sequences pairs of length $N = 2^n$ can be constructed, whose cross-correlation is $1 + 2^{[(n+2)/2]}$ [41]. Although this CC can get very small for long codes, similar to the auto-correlation properties, the cross-correlation properties degrade significantly for

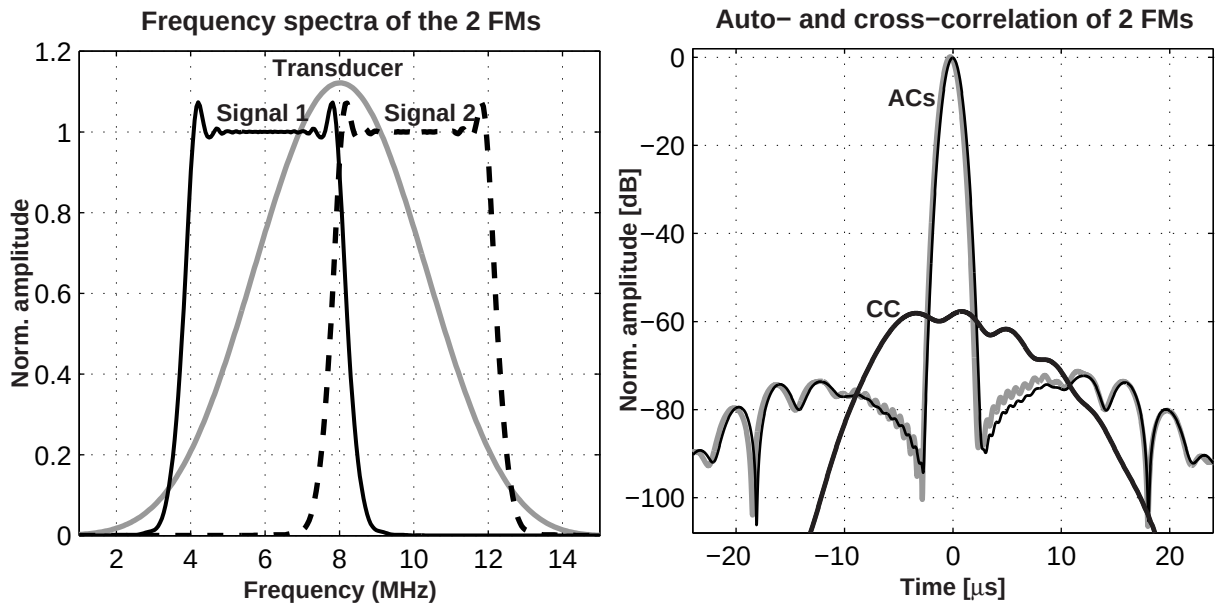

Figure 8.5: Frequency spectra and correlation properties of two tapered FM signals with the same sweeping bandwidth and frequency division.

short code lengths. For instance, two m-sequences of length 127 will have a cross-talk of 14.6 dB.

One of the advantages of binary sequences over the linear FM waveform is the big number of sequences for a given code length with good CC properties. An excellent discussion on CC properties of binary phase codes, particularly m-sequences, Gold and Kasami sequences is given in [41]. For codes of length N and family size (number of codes) M, the optimal theoretical minimum of the periodic CC peak is:

$$R_{xy\max} \geq N\sqrt{\frac{M-1}{MN-1}} \tag{8.5}$$

There is a vast amount of literature written on construction methods for sets of binary sequences that are as close to this limit as possible. Some families of sequences approach the limit of (8.5) for large values of N. However, for short lengths, CC values are significantly degraded. For instance, Gold sequences of length $N = 2^n - 1$ offer large families with a family size $M = 2^n + 1$. For Gold sequences of length 63 (n=6), their cross-correlation values are only 54% from the optimal bound of (8.5), resulting in 63 signals with maximum CC peak only 6 dB below the AC value [60]. For Kasami sequences with smaller family size $M = 2^{n/2}$, the CC peak for N=63 is 21% from optimal, resulting in 6 sequences with maximum CC peak of about 9 to 10 dB below the AC value. In practical terms that means, that in medical imaging, where code length has to be kept short, the cross-talk among phase-coded signals will be at the 5-15 dB range, depending on the number of signals and code family.

Aditionally, from the discussion in Chapter 5 it follows that the small code length allowed in

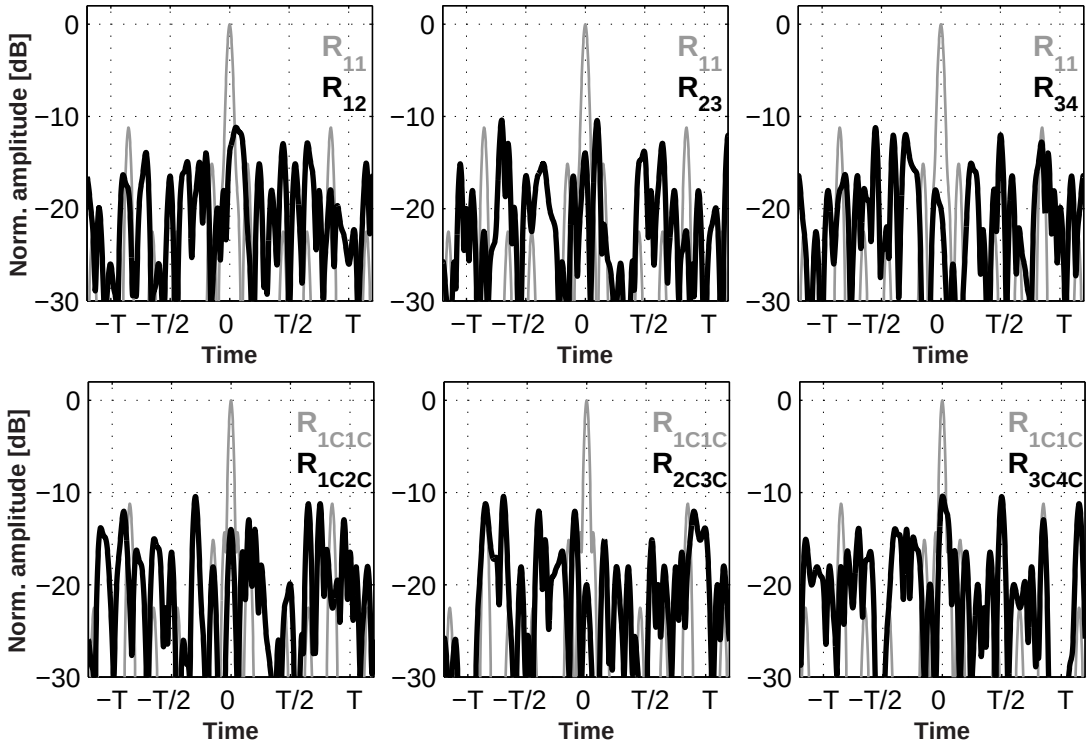

Figure 8.6: Cross-correlation functions between four sequences and between their four complementary sequences taken from four Golay pairs. The maxima in the cross-correlation functions are all below -10 dB relatively to the auto-correlation between any of the codes.

ultrasound imaging, makes binary codes, with the possible exception of complementary Golay codes inappropriate in terms of resolution. The limiting factor is the auto-correlation properties rather than the cross-correlation properties of binary codes. Therefore, only the cross-correlation properties of Golay codes will be further discussed.

The number of different Golay sequences of length 2^N is $(2^N/2)N!$. The number of different PN sequences is determined by the number of different generating polynomial, which is $E(2^N - 1)/N$, where E is the Euler function. For instance, there are 23040 Golay sequences and 378 PN sequences of length 64. For Golay codes of other lengths, the number of Golay codes is smaller, but still at the same order as the PN sequences [43]. Properties and construction methods of Golay codes have been described by Dokovic [42] and Budisin [43]. There are essentially two algorithms for generating Golay pairs: a) a non-recursive method originally introduced by Golay in the sixties, and b) a recursive method. Golay pairs of higher length MN can be constructed by a multiplication method of two Golay pairs, one of length M and one of length N. Dokovic [42] has reported a list with all Golay pairs up to length 40, which can be constructed with this method.

From his data, we have evaluated the cross-correlation between 220 different Golay codes, as well as the cross-correlation between their complementary codes. The maximum amplitude of the

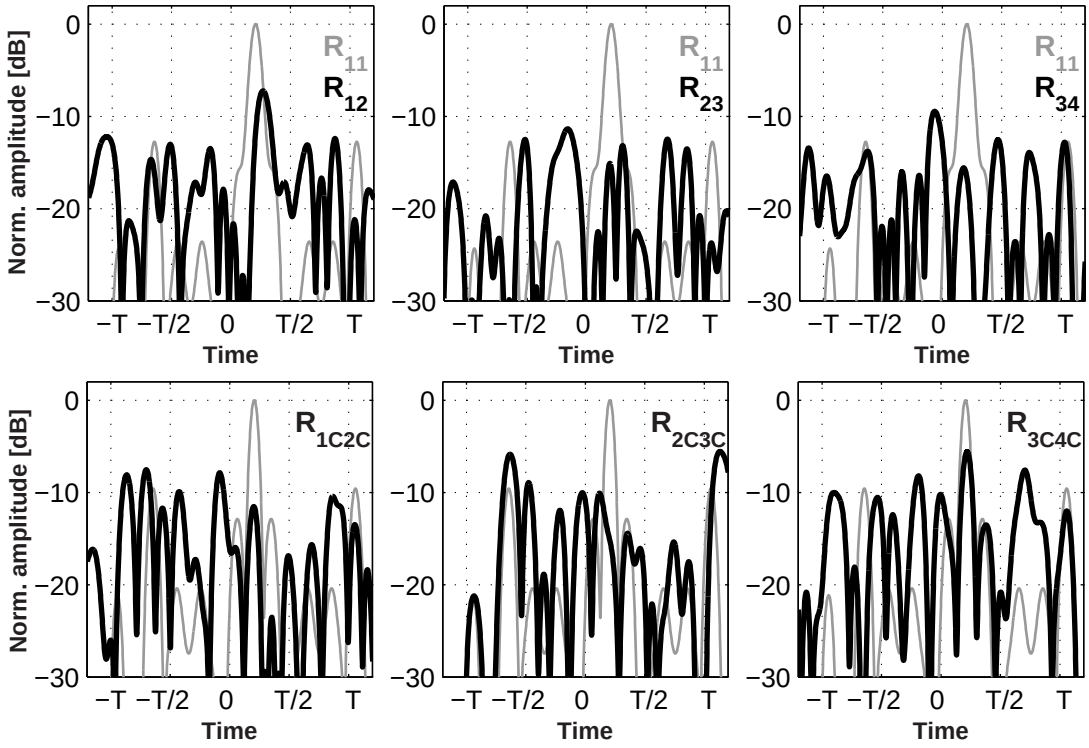

Figure 8.7: Cross-correlation functions between four sequences and between their four complementary sequences when all signals are convolved with the pulse-echo transducer impulse response. The maxima in the cross-correlation functions vary from -5.6 to -11.4 dB relatively to the auto-correlation peak.

cross-correlation among various codes is between 3.5 and 11.2 dB. We have chosen heuristically the following 4 Golay pairs, whose cross-correlation properties between them are below 10 dB:

G_1 = {1,1,1,1,-1,-1,-1,-1,-1,-1,-1,-1,-1,-1,-1,1,1,-1,-1,-1,-1,-1,1,-1,1,1,-1,1,-1,-1,1,1,-1,-1,1,1,-1,-1,1,-1,1}
G_{1C} = {1,1,1,1,-1,-1,1,1,-1,-1,1,1,-1,-1,-1,-1,1,1,1,1,1,-1,1,-1,1,1,-1,-1,1,-1,1,-1,1,-1,1,-1,1,-1,1,1,-1,-1,1,-1}
G_2 = {1,1,-1,-1,-1,-1,-1,1,1,-1,-1,-1,-1,1,-1,1,-1,1,1,1,-1,-1,1,1,-1,-1,1,-1,1,-1,1,1,-1,-1,1,-1,-1,-1,-1,-1,1,1}
G_{2C} = {1,1,-1,-1,-1,-1,-1,1,1,-1,-1,1,1,-1,1,-1,1,-1,-1,1,1,1,1,1,-1,-1,1,-1,1,-1,1,1,1,1,1,-1,1,1,1,1,1,-1,-1}
G_3 = {1,1,-1,1,1,-1,-1,1,1,-1,-1,1,1,1,1,1,1,1,-1,1,1,1,-1,1,1,-1,1,1,1,1,-1,1,1,-1,1,-1,1,-1,-1,1,1,1,-1,-1,-1,-1,1}
G_{3C} = {1,1,-1,1,-1,-1,1,-1,-1,1,1,1,1,1,1,1,-1,1,1,1,-1,-1,-1,1,-1,-1,-1,1,-1,1,-1,1,1,-1,-1,-1,1,1,1,1,1,-1}
G_4 = {1,1,-1,-1,1,1,1,-1,1,1,-1,-1,1,-1,-1,1,-1,1,1,1,1,1,1,-1,-1,1,1,1,-1,-1,-1,1,-1,1,1,1,-1,-1,1,-1,1,1,1,1,1,1}
G_{4C} = {1,1,1,1,1,-1,1,1,-1,-1,1,1,-1,1,-1,-1,-1,1,1,1,-1,-1,-1,-1,-1,-1,-1,1,-1,1,1,1,-1,1,-1,1,-1,1,-1,-1,-1,1,1,1,-1,-1}

Fig. 8.6 shows the cross-correlation functions between those four sequences and between their four complementary sequences. The highest cross-talk between any two signals or their complementaries is 10.4 dB below the auto-correlation peak.

However, cross-talk between Golay codes increases when all signals are convolved with a simulated pulse-echo transducer impulse response. The worst case is for the CC between the comple-

mentary Golay sequences G_{2C} and G_{3C} and between G_{3C} and G_{4C}, where the maximum reaches -5.6 dB (Fig. 8.7)

CHAPTER

NINE

Fast coded array imaging

This chapter investigates the possibility of increasing the frame rate in ultrasound imaging by using coded excitation. Linear-array imaging and synthetic aperture imaging will be considered, and the trade-offs between frame rate, image quality and SNR will be discussed.

9.1 Linear array coded imaging

From the discussion in Chapter 8, it follows that the peak value of the cross-correlation among binary codes of length N can not be less than about $\sqrt{2N-2}$. The exact lower limit depends on the choice of the code, with the m-sequences and Kasami sequences giving the best cross-correlation properties [41]. For practical codes of length 30-100 which can be used in ultrasound imaging, this corresponds to a cross-talk level of at best 15 dB below the auto-correlation peak for the case of two codes. The dynamic range of ultrasound B-mode images is at least 45 dB for most scanners, and, thus, such interference level will significantly degrade the image quality. For the FM signals, where the signal diversity is more limited, the best cross-talk level that can be expected for the case of only two signals in the set that share the same bandwidth is 12-20 dB. If more than two signals are needed, the cross-correlation among them will be in the range 6-10 dB.

Since the cross-correlation properties alone do not suffice for high quality fast imaging, imaging strategies have to be sought to minimize acoustic interference among the transmitted signals. If, for instance, these quasi-orthogonal excitation signals are focused in different image directions, so that this additional 25-35 dB interference reduction is achieved from the beam formation process, an increase in frame rate would be possible.

It is, therefore, first necessary to investigate the effect of focusing on cross-talk reduction, i.e. the lateral resolution of a linear-array imaging system. This has been done with the simulation program Field II [56]. The simulation parameters for the results presented in this section are the ones listed in Table 9.1, unless otherwise is stated. A 128-element transducer with a nominal frequency of 4 MHz is used.

Fig. 9.1 shows the beam formation process in which a sub-aperture of a linear array transducer

Table 9.1: Simulation parameters for linear array imaging

SIMULATION PARAMETER	VALUE
Central frequency	4 MHz
Element pitch	λ
Element height	5 mm
Number of elements	128
Transmit aperture size	22 elements
Transmit apodization	Hanning
Transmit focus	Fixed at 70 mm
Receive aperture size	48 elements
Receive apodization	None (rectangular)
Receive focus	Focal zones every 5 mm

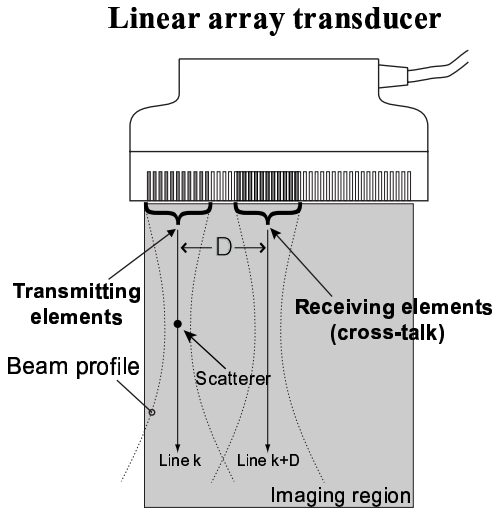

Figure 9.1: Illustrated method for the evaluation of the lateral resolution in linear array imaging. A group of elements transmits a focused beam along line k, and two lines k and $k+D$ are beamformed using two receive sub-apertures.

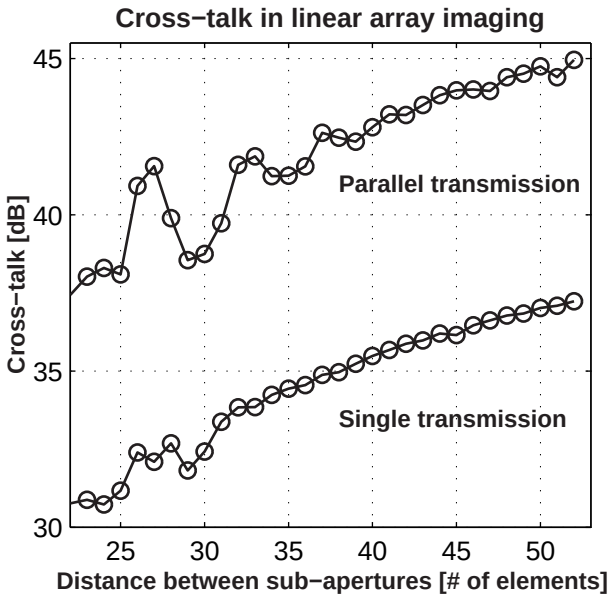

Figure 9.2: Cross-talk for single and parallel transmission in linear array imaging. The first case is the conventional imaging, where a focused beam is transmitted by a sub-aperture and the amplitude of the echoes from a moving receive sub-aperture is measured. In the second case, both sub-apertures transmit simultaneously using two different FM-coded signals with different slopes.

transmits a focused beam along line k. Line k is formed by receiving in this case with the same transmitting sub-aperture. Assuming a medium with only one scatterer positioned at the focal depth along line k, echo signals should ideally be present only in the beamformed data for line k. However, due to the finite sub-aperture sizes, the limited lateral resolution of the pulse-echo imaging system causes weaker echoes to be also present in the beamformed data for line $k+D$.

Fig. 9.2 shows the ratio in decibels between the peak value of the beamformed data from line k and the peak of the beamformed data from line $k+D$ as a function of the distance D between the two receive apertures. The curve labeled single transmission corresponds to the cross-talk when only the left sub-aperture transmits. In case of single transmission, it is immaterial in the cross-talk whether pulse or coded excitation is used. This cross-talk depends on the number of elements on the transmit and receive apertures as well as on the apodization functions of both transmit and receive apertures. For instance, the cross-talk reduces when the transmitting sub-aperture is apodized, since the transmitted beam becomes more focused. Thus, the simulation results of Fig. 9.2 are only indicative of the order of the expected cross-talk under realistic imaging conditions.

The second graph of Fig. 9.2 labeled parallel transmission shows the cross-talk when both sub-apertures transmit different coded signals. In this simulation two linear FM signals with the same bandwidth and opposite FM slopes are used. Recall from Chapter 8 that the cross-talk between two such signals is about -12 dB. The received rf-data from each of the two receive apertures are correlated with the corresponding compression filter and subsequently are beamformed to yield

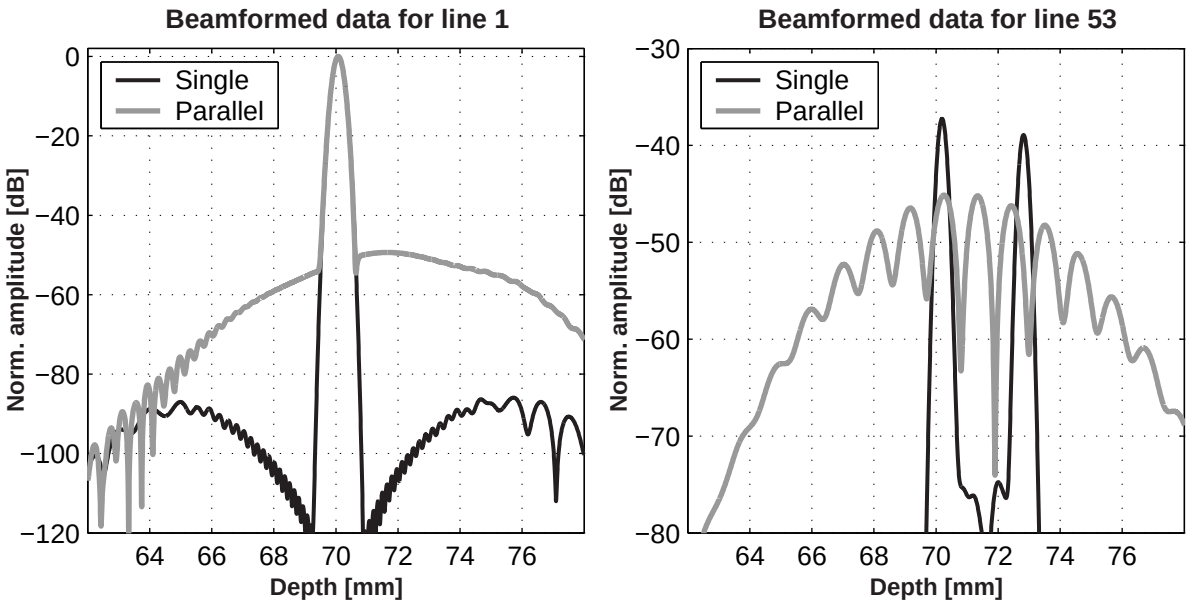

Figure 9.3: The effect of simultaneous transmission of two beams in axial and lateral resolution. For parallel transmission of two beams, the cross-talk in the second channel reduces, but the axial resolution of the measurement in the first channel becomes limited by the cross-talk.

lines k and $k+D$ as previously. Then the cross-talk is plotted as a function of the distance D between the sub-apertures. The cross-talk is reduced by about 7.5 dB, which is less than the cross-correlation between the two signals. The reason for that is that the received echoes from the second sub-aperture is a summation of two terms: i) the echoes resulting from the overlap between the lateral sidelobes of the first transmitted beam with the sidelobes of the second received beam, cross-correlated with the second code and ii) the auto-correlation of the echoes due to the sidelobe overlap between the second transmitted and received beam.

The consequences of simultaneous transmission of two coded beams are twofold: the first is the reduction in cross-talk for the corresponding lines that are beamformed simultaneously as shown in Fig. 9.2. This interference will appear as ghost echoes in areas of the image where no scatterers actually exist. What this figure does not show, however, is how the axial resolution is affected. The axial sidelobes will result in shadows around strong scatterers, that can mask weaker scatterers along this direction. Fig. 9.3 shows the effect of parallel transmission from two beams. The centers of the two transmitted beams are spaced as far away as possible (53 elements apart). This corresponds to the far right points of Fig. 9.2. The axial resolution of the point scatterer located at the focal point of the first transmitted beam degrades due to the cross-talk of the second transmitted beam. The cross-talk in the second beam reduces from 37.2 to 45.1 dB as shown also in Fig. 9.2.

Inspecting the simulation results just presented, the following conclusion can be drawn. For the used simulation parameters, the best case in cross-talk without any coding is 37.2 dB, occurring when the transmit-receive sub-apertures are as far apart as possible (most right points in Fig. 9.2). If both sub-apertures transmit simultaneously, there will be an increase in frame rate by a factor

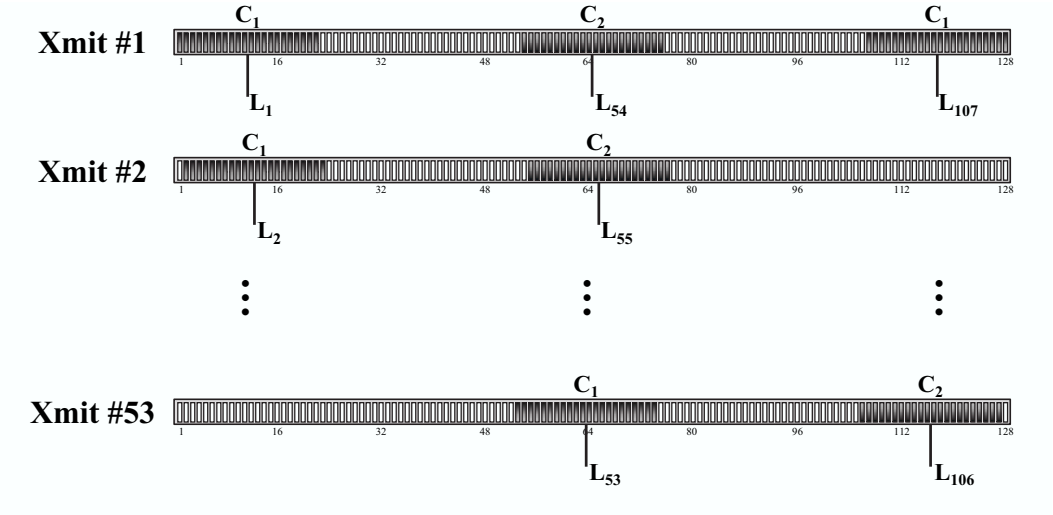

Figure 9.4: Firing sequence in coded linear array imaging with double frame rate. In the first transmit event, three lines are formed simultaneously, while in all other transmit events, two beams are formed. Two FM signals of different slope are used for parallel transmission.

of 2, but there will also be sidelobes of 37.2 dB both axially and laterally. Assuming a desired dynamic range of 45 dB in the ultrasound image, **the frame rate in linear array imaging can be doubled without any degradation in image quality, by using two coded sequences that have a cross-correlation of at least 11 dB.** The firing sequence for linear array coded imaging with double frame rate compared to conventional linear array imaging is shown in Fig. 9.4. Two beams are transmitted for every transmit event, while three beams are sent in the first transmit event. Fig. 9.5b shows that linear array imaging with 2 parallel FM-coded beams yields good resolution images with a 45 dB dynamic range. The firing sequence is that shown in Fig. 9.4. Fig. 9.5a shows the conventional linear array image for comparison. Binary phase codes or the linear FM signals with frequency division discussed in the previous chapter can also be used. In the latter case, all sidelobes shown in Fig. 9.3 will be eliminated down to -90 dB, with a 80% widening of the axial mainlobe.

9.2 Other firing and coding strategies

Further increase in frame rate is possible by reducing the distance between the transmitting sub-apertures. Fig. 9.6 shows an alternative firing sequence, where three or four beams are sent in parallel.

Fig. 9.7 shows the resulting simulated image when two FM signals with opposite slopes are used in the firing order shown in Fig. 9.6.

The sidelobe performance of this imaging scheme can be assessed from Fig. 9.2. In the first transmit event, the cross-talk between the parallel beamformed line pairs (1,23), (23,45), (45,67)

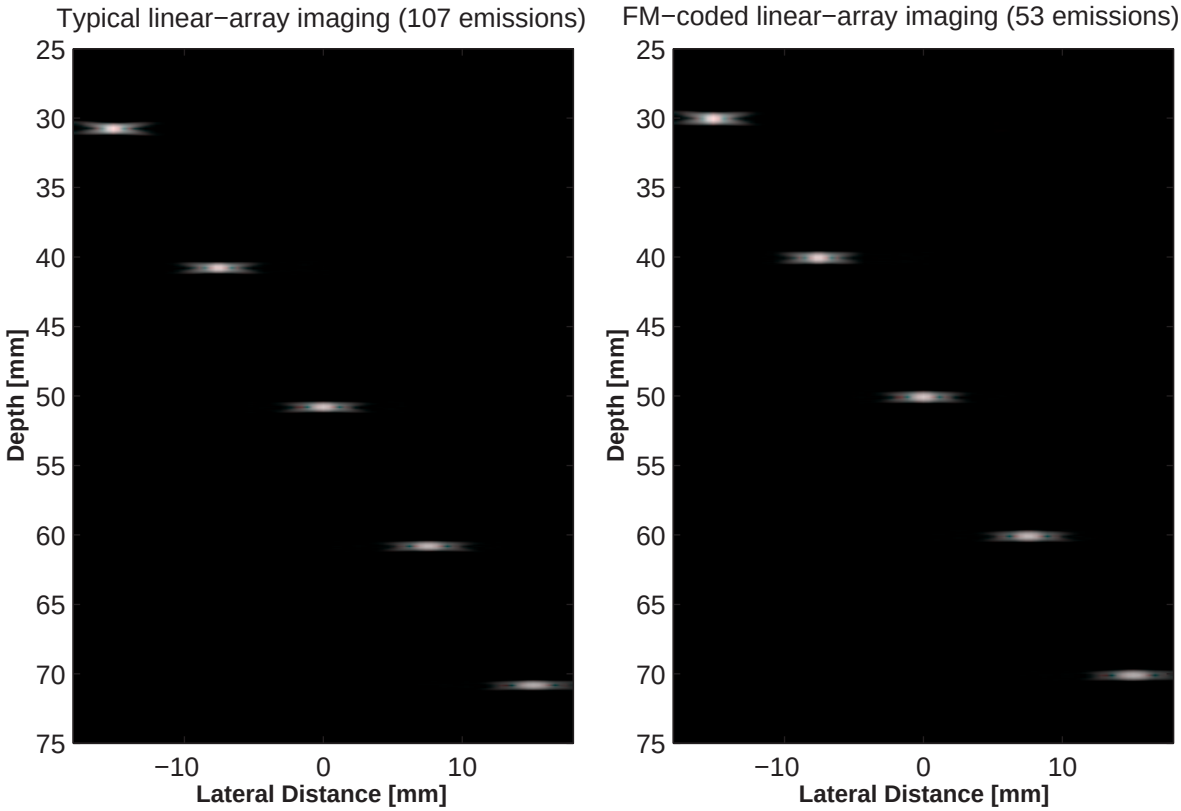

Figure 9.5: Conventional linear array imaging (left) and linear array imaging with double frame rate using two parallel FM-coded beams. The dynamic range of both simulated images is 45 dB.

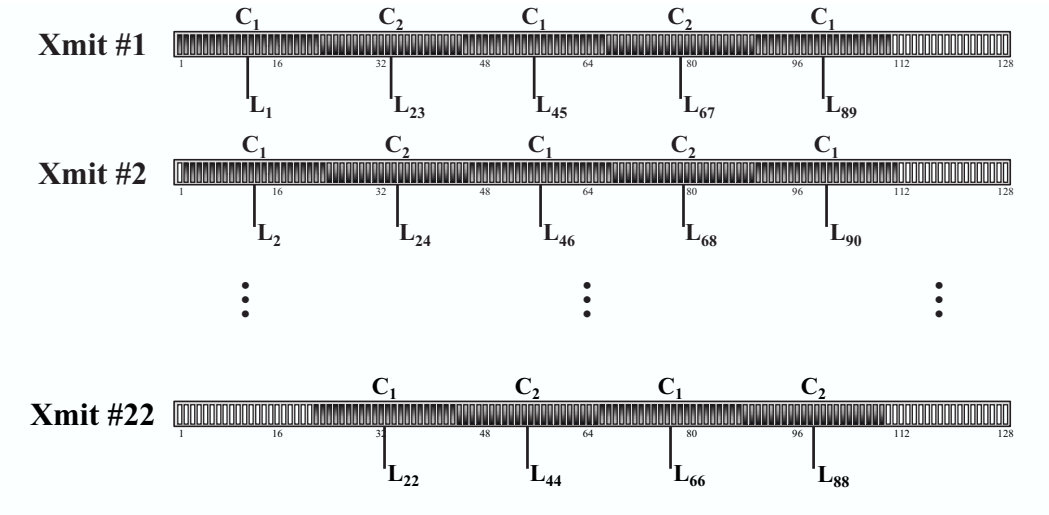

Figure 9.6: Alternative firing sequence in coded linear array imaging where three or four beams are sent in parallel. The number of transmit events reduces from 107 (conventional imaging) to 22.

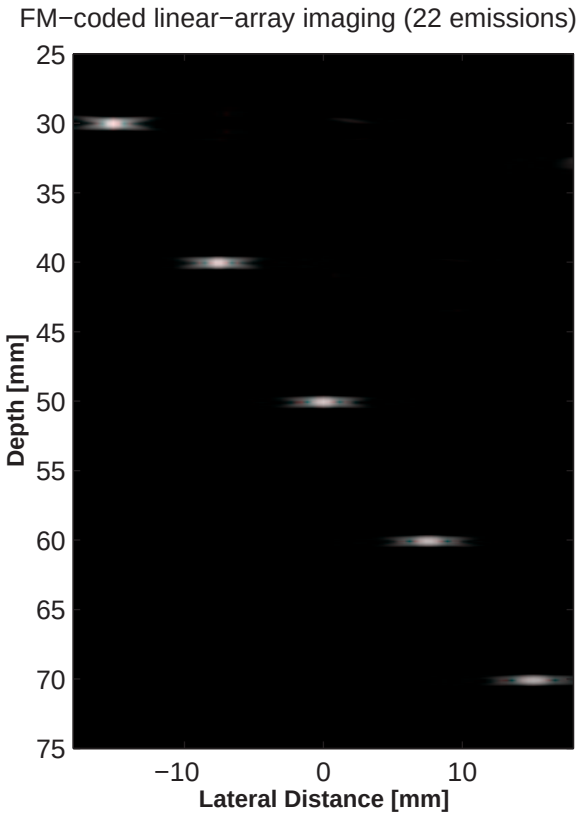

Figure 9.7: Simulated image of fast FM-coded linear array imaging using the firing scheme of Fig. 9.6. The number of transmit events is almost 5 times less than in conventional imaging. The dynamic range of the image is 45 dB.

and (67,89) is 37.5 dB (upper graph of Fig. 9.2 for $D = 22$). This can be reduced by using a set of coded signals with lower cross-correlation, such as binary codes or the FM signals with frequency division. The cross-talk between the line pairs (1,45), (45,89) and (23,67) is 36.2 dB (lower graph of Fig. 9.2 for $D = 44$). This is the acoustic cross-talk between lines that use the same transmitted code. The ghost echoes that are visible at a depth of 35 mm in Fig. 9.7 manifest the presence of these sidelobes.

It is therefore clear that fast imaging with the firing scheme of Fig. 9.6 using only two coded signals can not eliminate the deleterious effects of acoustic interference. If more than two coded sequences are to be employed, the FM coding has to be abandoned, since this family of signals can not offer such diversity.

One possibility is the transmission of four Golay pairs one for each of the parallel beams of Fig. 9.6. The cross-correlation between the four Golay-coded signals as well as the cross-correlation between their complementary codes is around 9 to 12 dB, depending on the chosen Golay signals. This ensures low lateral sidelobes. The drawback of this method is that two emissions are required for every line, one using the Golay codes and one using their complementary Golay codes. Therefore, the number of emissions in this case is 44.

9.3 Synthetic transmit aperture (STA) imaging

Phased array imaging, currently used in ultrasound scanners, involves transmission of pulses from the entire transmit aperture, which are relatively delayed to form a focused beam along a given direction. The echoes received by all elements are beamformed to yield the image points along this direction. In linear array imaging focused beams along a line are transmitted and received by the same sub-aperture. There are two main limitations in both methods: i) the acquisition time is proportional to the number of lines in the image, ii) the image has a fixed transmit focus.

Synthetic transmit aperture (STA) imaging can overcome both these problems. In STA imaging, one element transmits a pulse and all elements receive the echoes. Since each transmission is a spherical wave insonifying the entire imaging region, receive beamforming for all lines can yield a whole low-energy image for every transmission. Then the next element is excited and this is repeated until the whole transmit aperture is synthesized. The transmit sequence for sparse STA imaging with four emissions is shown in 9.9.

The final synthetic image is the coherent sum of all beamformed images [61]. In this way, the final image is optimally focused both in transmit and receive.

The frame rate in synthetic aperture imaging does not depend on the number of scan lines to be formed as in conventional imaging, but on the number of transmit elements. If only two to five elements are used for transmitting, a frame rate on the order of 1,000 frames/s is possible, which will pave the way to 3-D imaging. The use of such small number of transmit elements is necessitated by the problem of artifacts caused by tissue or transducer motion. In contrast to linear and phased array imaging, where each line is formed from a single transmission, in synthetic

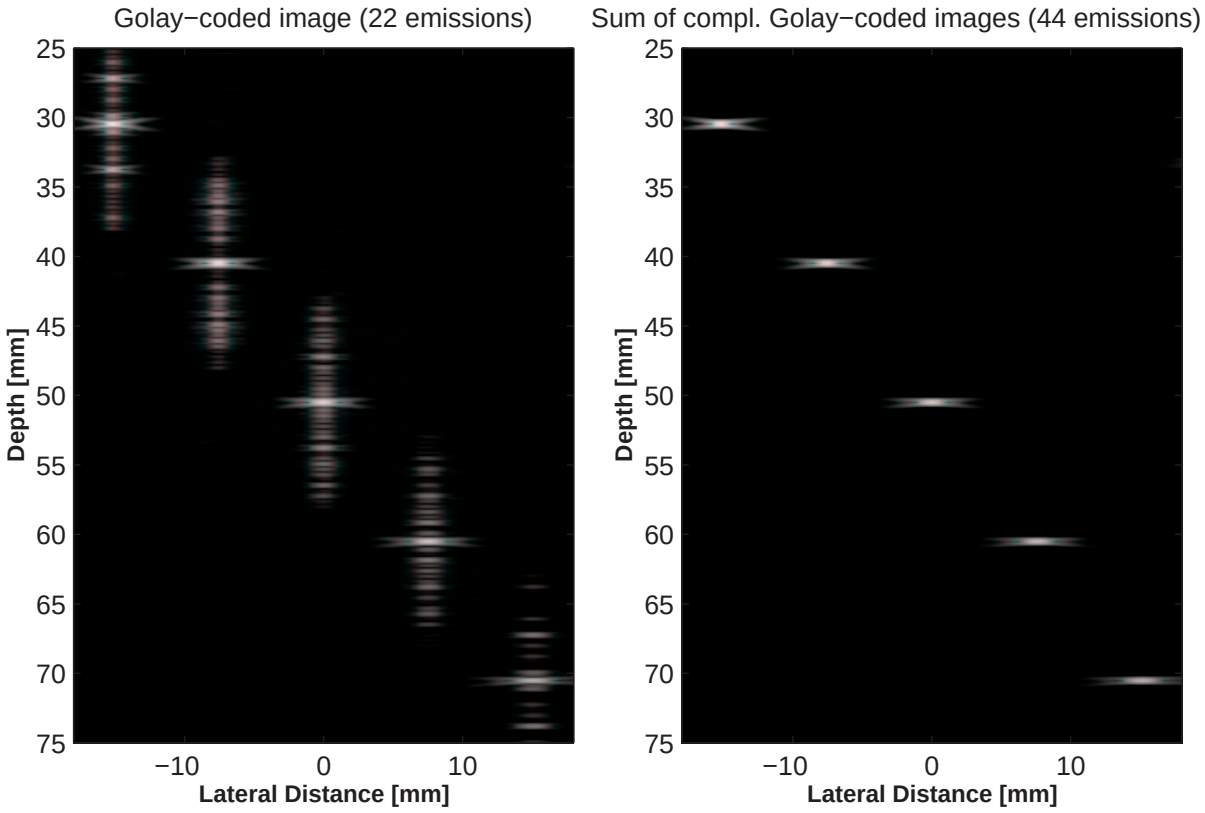

Figure 9.8: Simulated images of fast coded linear array imaging employing four Golay pairs. The image on the left is one of the two images using Golay codes. The image on the right is the summation of the two complementary Golay-coded images. The dynamic range of both images is 45 dB.

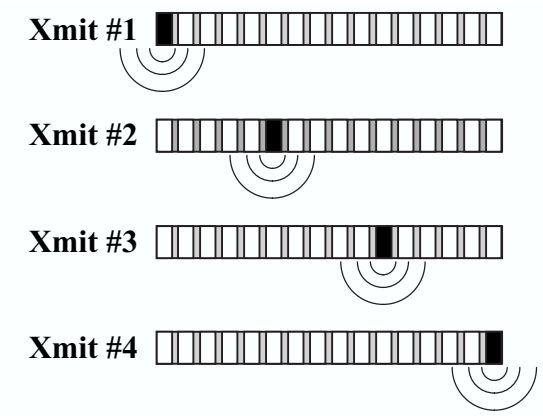

Figure 9.9: Transmitting succession scheme for sparse synthetic transmit aperture imaging using four emissions. One element sends out a spherical wave for every transmit event and all elements receive the echoes. All beams are formed simultaneously for every transmit event.

aperture imaging each line of the final image uses data from all transmit events, which sets phase coherence requirements. Using 5 transmissions with an 8 MHz transducer and a corresponding wavelength of about 190μm, tissue speed greater than 14 mm/s may cause motion phase artifacts.

Such sparse transmit aperture in STA imaging can yield a high frame rate, no motion artifacts, and potentially optimally-focused images, but rises two new issues. The first is the resulting grating lobes in the radiation pattern of the array, since the necessitated sparse transmit aperture violates the half-wavelength element spacing. This problem has been discussed in the literature [61, 62] and will not be addressed here. It can be solved using the effective aperture concept. Briefly, the convolution of the transmit with the receive aperture for each transmission is the effective sub-aperture. The two-way radiation pattern of the synthetic image is the Fourier transform (at least in the far field) of the sum of the effective sub-apertures of all transmit events. Therefore, by proper apodization of the receive aperture for each transmit event, a uniform $\lambda/2$ spacing effective aperture can be created, even in the case of only 2 transmit events [62].

The second problem of sparse STA imaging compared to conventional array imaging is the low signal-to-noise ratio (SNR). Utilization of coded excitation can make it possible to increase the frame rate without sacrificing the SNR. The coding methods for fast STA imaging with high SNR will be described in the following sections.[1]

[1]Part of this work has been submitted for publication in *Ultrasonics*, 2002.

9.4 Literature review on SNR improvement methods in STA imaging

It has been reported [63] that the SNR for a non-sparse STA imaging system, which uses all transmit elements N, is $20\log(\sqrt{N})$ less than the SNR of an equivalent phased-array imaging system. This is because the power transmitted by all elements in phased-array imaging is at the focal point N times higher than the power transmitted by 1 element as in STA imaging. Then, for a 64-element system, there should be a degradation in SNR of $10\log 64 = 18$ dB. For a STA system with 4 transmissions, the expected SNR degradation will then be at the order of 30 dB.

There have been three methods reported in the literature for increasing the SNR in STA imaging. In the first [63], transmit defocusing delays are applied to a group of elements, in order to obtain a similar one-way radiation pattern as a single transmitting element at the center of the group. In this case, the mean intensity of the emulated spherical waves increases by n (and not n^2 due to defocusing), as demonstrated experimentally by Karaman et. al. [63], where n is the number of simultaneously transmitted elements. For 11 elements an improvement in SNR of about $10\log 11 = 10.5$ dB should be expected.

In the second method [64], all N elements simultaneously transmit spherical waves with excitation pulses spatially encoded by the rows of a $N \times N$ matrix, such as the Hadamard matrix. The spatially-coded transmit sequence is shown in 9.10.

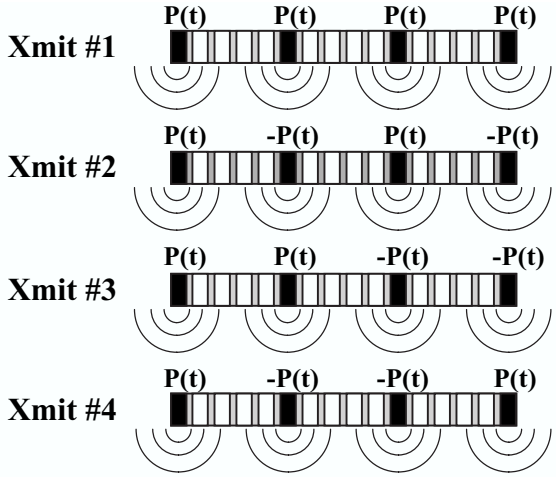

Figure 9.10: Transmitting succession scheme for sparse STA imaging using Hadamard spatial encoding. All active transmit elements send out spherical waves for every transmit event and all elements receive the echoes, which are decoded by the inverse matrix before beam formation.

After N spatially-encoded transmit events, the received data sets can be decoded, by the inverse matrix, yielding the equivalent N data sets as if each of the N elements was transmitting alone in every transmit event. The advantages of the Hadamard matrix over other orthogonal matrices are: i) it only consists of 1 and -1 (making coding implementation easy as a phase reversal of the

transmitting pulse), and ii) the inverse Hadamard matrix is a scaled version of itself. Hadamard spatial encoding increases SNR by $10\log M$ [64], where M is the number of simultaneous transmitting elements (or group of elements). For a 4×4 Hadamard matrix considered here, there is an improvement of 6 dB.

The third reported method combines coded excitation with spatial Hadamard coding [65, 12] . For four transmit events, it uses two pairs of complementary codes, while the four codes are also mutually orthogonal. The decoding process is based on the Hadamard matrix of half length, assisted by the cancellations of the sidelobes in the sum of complementary auto-correlations and cancellations of the sum of cross-correlations among mutual orthogonal codes. The improvement in the SNR is $10\log(M \cdot TB)$, where TB is the time-bandwidth product of the codes, which is equal to the length of the code L for binary codes. Using four transmit events and codes of length 32, as reported in [65], there is a $10\log 4 \cdot 32 = 21$ dB improvement.

9.5 Proposed STA coded imaging using Hadamard and FM space-time encoding

The encoding scheme proposed here is a modification of the aforementioned method. It is apparent that the SNR can be significantly increased if long coded sequences are used instead of short pulses. It has been previously shown [66], that for the highly attenuating medium of medical ultrasound the linear FM waveform is the most appropriate coded waveform, as long as the pulse compression mechanism ensures low range sidelobes and good axial resolution.

A 64-element linear array with $\lambda/2$ element spacing at 4 MHz has been used in the simulations. The entire array was used in receive for all transmit events with the same Hanning window for apodization of the received aperture. Such apodization does not give the optimal effective aperture, and therefore higher lateral sidelobes compared to phased-array imaging are expected. For each of the 4 transmit events, the elements 1, 22, 43, 64 transmit the same tapered linear FM waveform. The design of the tapered FM signal and the weighted compression filter, which are used, have been described in [20]. The received signal, which has been convolved with the two-way transducer impulse response has an effective time-bandwidth product of 130. The compressed output from this compression scheme has a theoretical -20 dB mainlobe width of 1.48 λ and very low axial sidelobes (mainlobe-to-peak sidelobe ratio of -88 dB).

Upon reception, each of the 4 rf-data sets is decoded by the Hadamard matrix. The decoded rf-data are subsequently compressed using the same compression filter in all channels. The order of decoding and compression was immaterial in the simulations, however in real images decoding might be a lot more sensitive than compression and should precede. Finally, the decoded compressed rf-data are beamformed for all directions, yielding a low-energy image. Fig. 9.11 shows simulated encoded STA images before and after compression obtained with the simulation program Field II [56]. Fig. 9.12 shows that the resolution both laterally and axially remains practically the same, when Hadamard spatial encoding combined with FM time encoding is used. Also notice that the matched filter aligns the rf-data axially at the center of each scatterer, in contrast to pulsed exci-

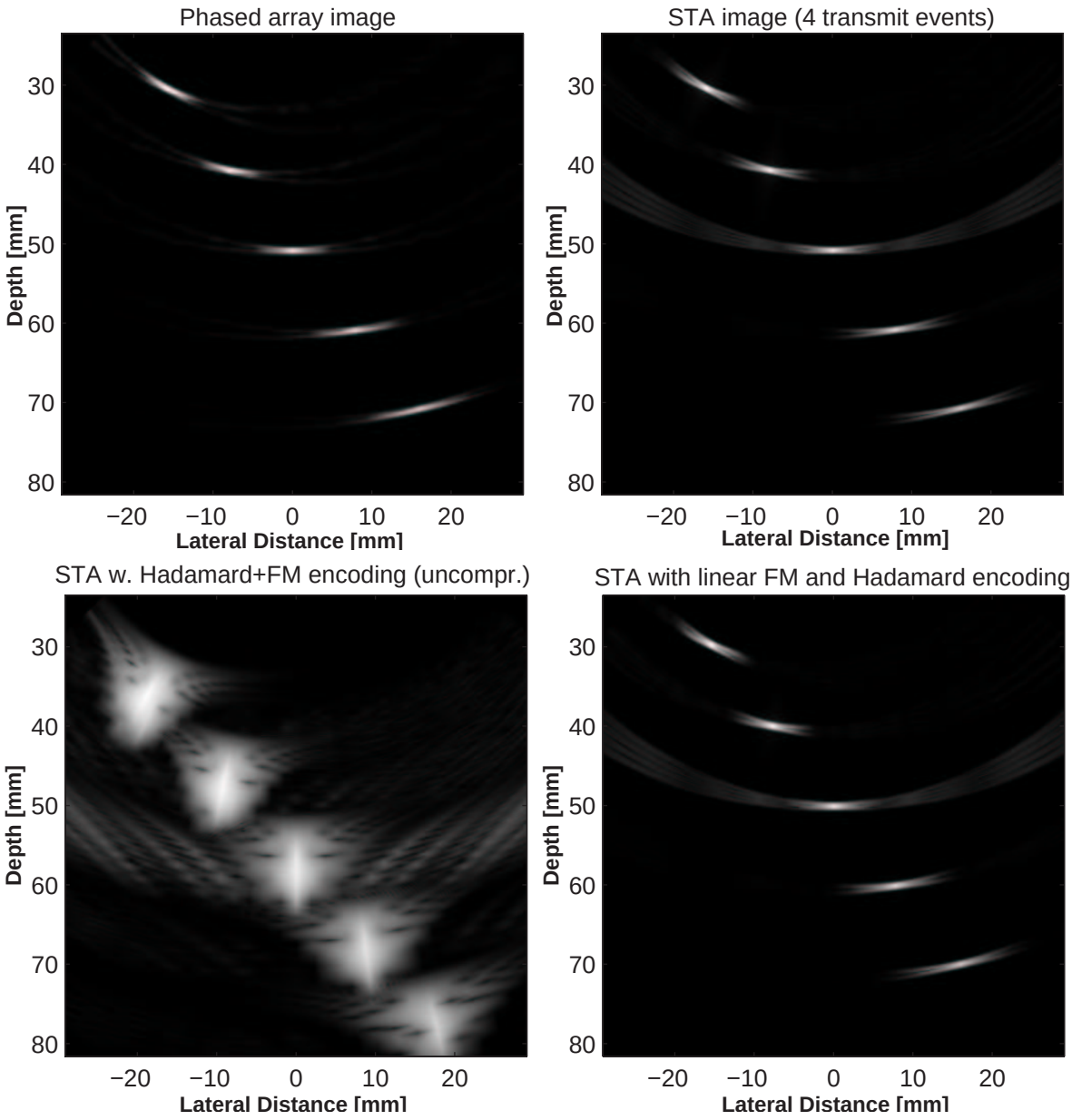

Figure 9.11: Simulated images of point targets. The first row of images is a conventional phased-array image and a typical uncoded STA image with 4 emissions. The second row shows coded STA images using Hadamard encoding and tapered linear FM signals, before (left) and after compression (right). The dynamic range of all images is 60 dB.

tation where each scatterer appears a little further than it actually is, due to convolution with the transducer impulse response.

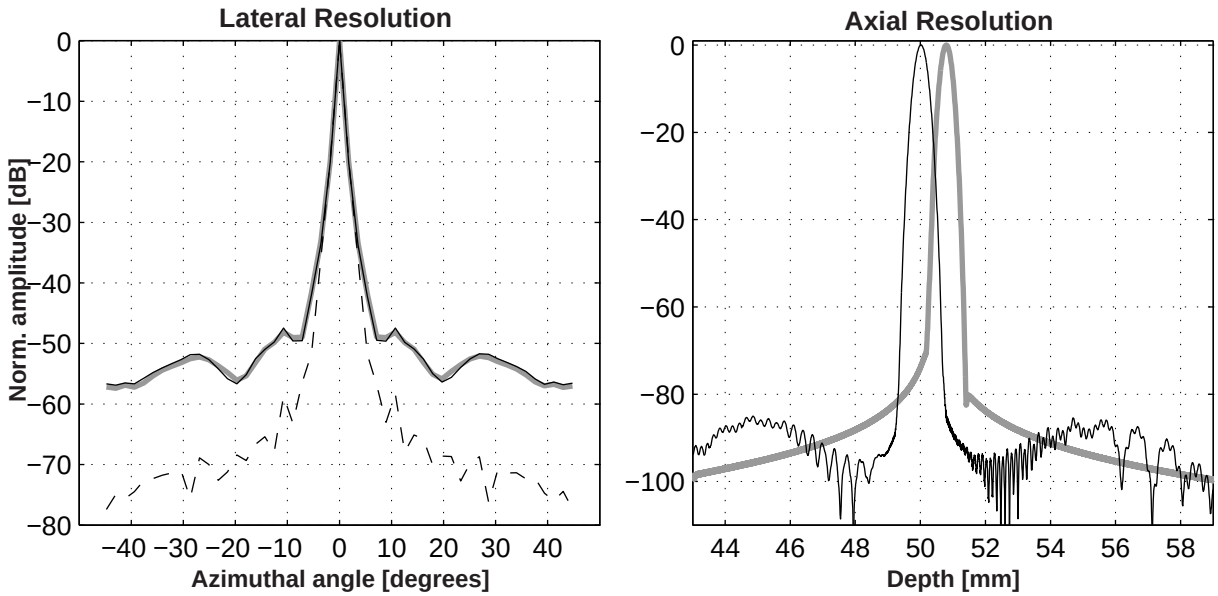

Figure 9.12: Lateral and axial resolution calculated from the simulated images at the point at depth 50 mm. The gray lines correspond to the typical STA image with 4 emissions, and the black lines to the STA image with the proposed Hadamard+FM encoding. The dotted line in the first plot shows the lateral resolution of the phased-array image.

This coding strategy yields the same frame rate and approximately the same SNR improvement as with the one reported in [65]. The advantage of the proposed method lies in the robustness of compression of linear FM signals, and the immunity this signal has in frequency shifts. On the other hand, the complementarity of Golay codes breaks down at large depths due to attenuation in tissues, resulting in image degradation. It has been shown [66] that with an attenuation of 0.7 dB/[MHz$\times$cm], the range sidelobes for a Golay pair increase up to -25 dB at a depth of 16 cm, in contrast to the proposed tapered FM signal, whose compression is very robust to attenuation. The temporal FM coding can also work independently in the absence of the spatial Hadamard coding, if the receiver has to be kept simpler. This is not the case for the complementary coding suggested in [65], where it is the Hadamard coding which cancels the sum of the cross-correlations between the mutual orthogonal codes. The approach of the Golay codes is also more sensitive to motion artifacts, since the cancellation of the sidelobes is based on two successive firings. Finally, in the proposed method, image quality can be traded off with doubling of the frame rate.

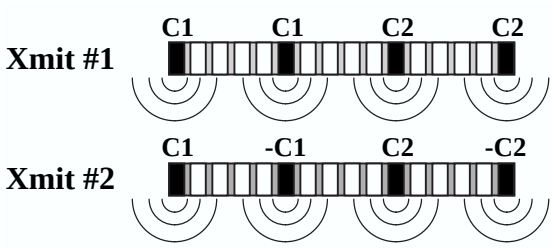

Figure 9.13: Transmitting succession scheme for fast sparse STA imaging using two orthogonal FM signals C_1 and C_2.

9.6 STA imaging with double frame rate using orthogonal FM signals

This can be done [67] by transmitting two linear FM signals from two pairs of transmitted elements, as shown in 9.13. If the two signals have low cross-correlation properties, a 2×2 Hadamard matrix can decode the four signals. In this way, two images can be beamformed for every transmit event and the frame rate is doubled at a cost of SNR and resolution. The two excitation signals are an up- and a down-chirp with frequency division. They have opposite FM slopes, equal sweeping bandwidth and different central frequencies. This design results in good auto- and cross-correlation properties discussed in Chapter 8. The axial sidelobes are less than -70 dB, and the cross-talk between the signals is around -60 dB. The axial resolution is inversely proportional to the bandwidth of the coded signals, and therefore with this method the axial resolution degrades by about 80%. A loss in SNR is expected since the transmitted energy is not centered at the transducer central frequency. The resulting STA image using only 2 emissions is shown in Fig. 9.14b. For comparison, the STA images with Hadamard and FM encoding using 4 emissions is shown at the left.

9.7 Evaluation of SNR in coded STA imaging

Fig. 9.15 shows simulated images of 5 point targets using the various coding schemes. White Gaussian noise was added to all the received channels before beamforming, to simulate electronic receiver noise. Equal driving voltages are used for all images. Note, however, that phased array imaging systems focus the pulses of all elements at the transmit focusing point. In STA imaging, only a single element transmits an unfocused wave. Therefore, for equal driving voltages, the transmitted power in STA imaging is much lower, and thus, the transmitted pulse amplitude can be increased more than an order of amplitude compared to conventional imaging without exceeding the intensity limitations [61]. It is, thus, possible to achieve up to 10 dB higher SNR than what is shown here, as long as the transducer elements can deliver such high power. The gain in SNR that can be achieved by using spatially encoded FM excitation is apparent from the images of Fig. 9.15. For quantitative evaluation, the SNR has been estimated from the central line of the images. The results are listed in Table 9.2. The entry for the phased array image is the maximum

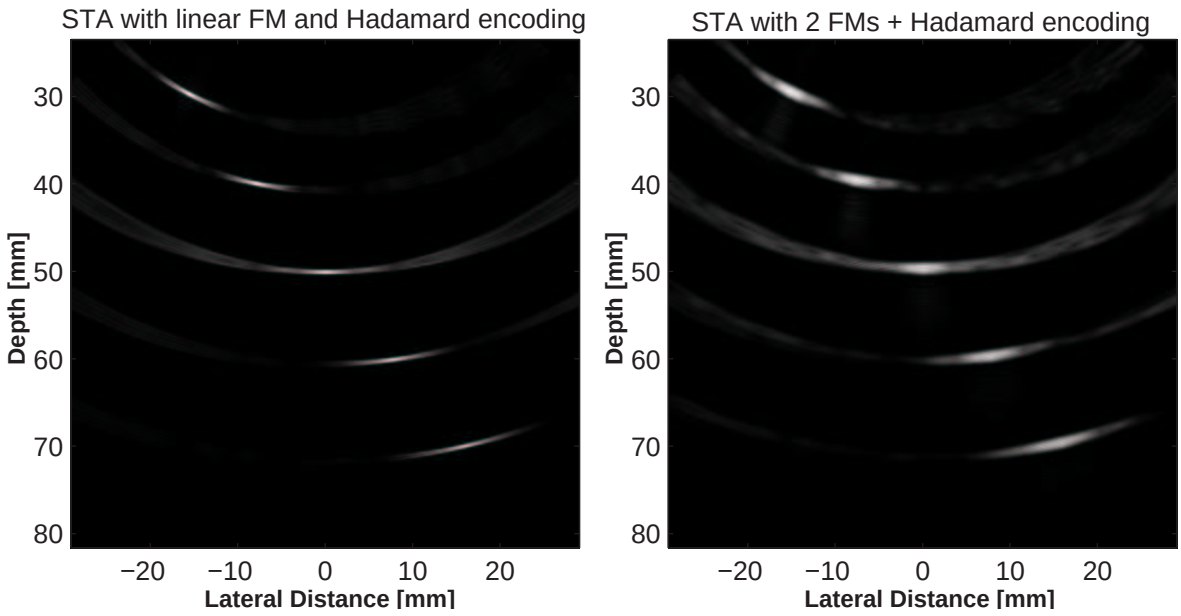

Figure 9.14: STA simulated image with double frame-rate (2 emissions) using Hadamard encoding and two orthogonal preweighted linear FM signals with frequency division (right). On the left, the coded STA image using 4 emissions for the same 8 MHz array transducer is shown for comparison. The dynamic range of the images is 60 dB.

SNR, calculated at the transmit focal point.

In the case of STA with 2 coded transmissions using orthogonal chirps, there is a trade off between low cross-correlation among the signals and loss in SNR. For a different design of the two codes than the one presented here, cross-correlation around -65 dB is possible with an additional loss of 5 dB in the SNR. In conclusion, the suggested coding scheme can yield synthetic aperture images with four transmissions, which can have the same SNR as that for phased-array imaging at the transmitted focal point, while retaining image resolution. In combination with the method of defocused transmit subapertures, this SNR can be achieved with only two transmissions of orthogonal preweighted linear FM signals.

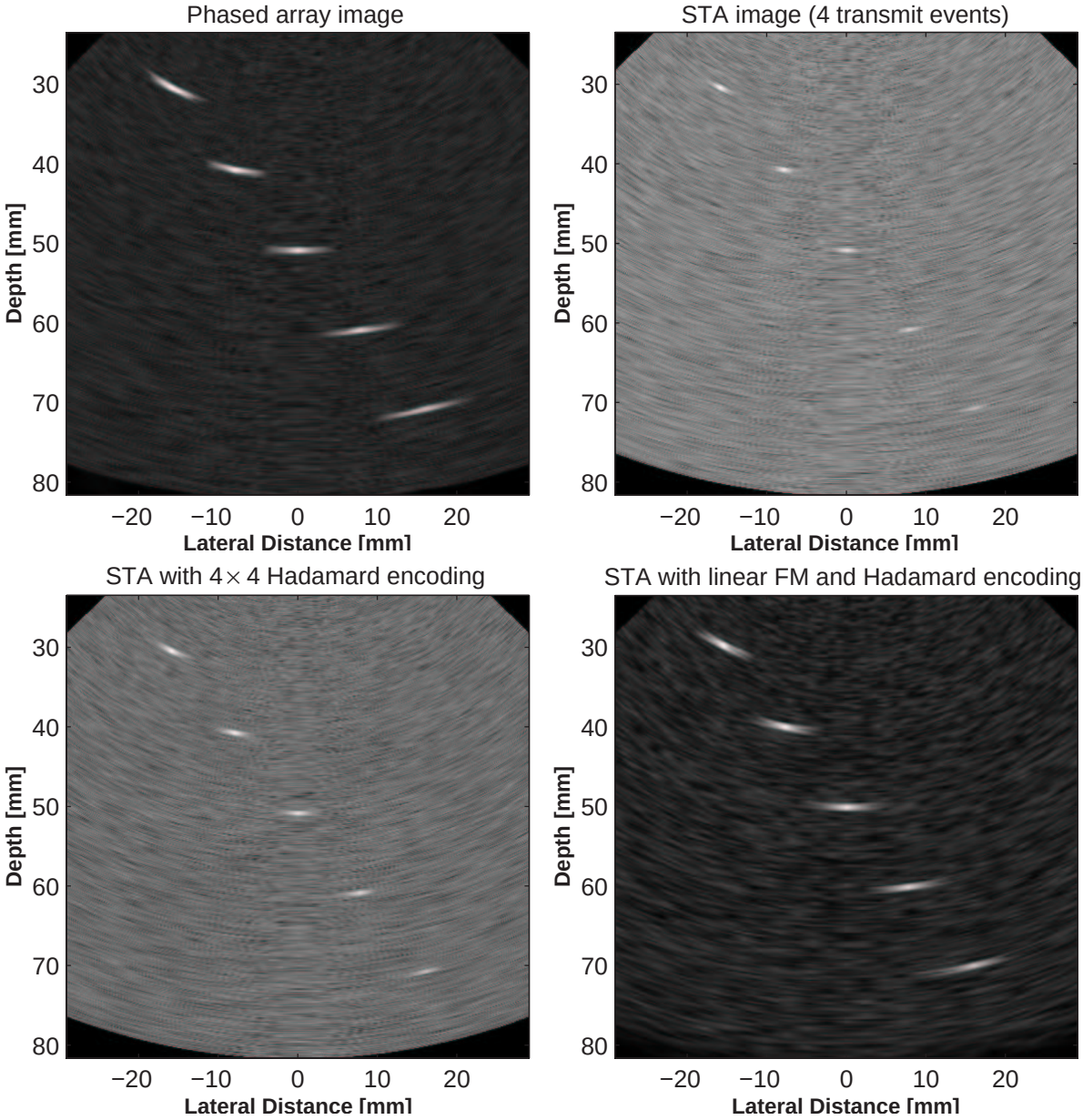

Figure 9.15: Simulated images of point targets with added noise, showing the improvement in SNR for the various coding schemes. The dynamic range of all images is 60 dB.

Table 9.2: Calculated SNR from the central line of the simulated images

IMAGING MODE	EMISSIONS	SNR (dB)	DEPENDENCE
Phased Array Imaging	51	53.09	$10\log(64\cdot 64)$
STA w. 1 xmit event	1	18.64	0
STA w. 64 xmit events	64	36.43	$10\log(64)$
STA w. 4 xmit events	4	24.34	$10\log(4)$
STA w. 4×4 Hadamard coding	4	30.33	$10\log(4\cdot 4)$
STA w. 4×4 Hadamard + linear FM	4	52.74	$10\log(4\cdot 4\cdot TB)$
STA w. 2×2 Hadamard + 2 linear FMs	2	47.56	$10\log(2\cdot 2\cdot(TB_1+TB_2))$

CHAPTER

TEN

Fast ultrasound imaging using pulse trains

The volume constrains of the ambiguity function which were discussed in Chapter 2, lead to the conclusion that thumbtack-type ambiguity surfaces will integrate their pedestal to a relatively high level. This will result in a poor signal-to-speckle ratio for a dense scattering medium. If this level is unacceptable for high quality imaging, alternative approaches have to be attempted. Instead of applying any coding that spreads the sidelobes, one could allow the spike ambiguities and try to resolve them. As it was shown in Chapter 2, the only high time-bandwidth product waveforms that can have clear areas in their ambiguity function are the pulse trains. The purpose of this chapter is to investigate if pulse trains can be used for fast ultrasound imaging.

Trains with linear frequency modulation from pulse to pulse are good candidates, since they have deep valleys parallel to the time axis. One approach is, thus, the transmission of two pulse trains with different positions of spike ambiguities. These trains can have different pulse duration or interpulse periods (a technique called staggering) or different coding. If the nulls of one waveform's matched-filter response are placed such as they fall at the ambiguity peaks of the matched-filter response of the other waveform, the ambiguities can be resolved. For effective cancellation, this has to be done on the envelope detected image data as a non-linear ”min” operator.

This chapter is organized as following: first coding possibilities of pulse trains are discussed. Then the ambiguity function of pulse trains is derived and the effect of coding and staggering are illustrated. Subsequently a new imaging method is presented, which utilizes pulse train excitation. With this method, a single pulse is applied to a transducer element and a pulse train is created acoustically at every imaging point. Simulated images of the method are shown and finally other embodiments of the proposed method are discussed.

10.1 Pulse trains

Repetition of pulses is an alternative method for constructing waveforms with large time-bandwidth products. If the same basic pulse is repeated, the resulting waveform is the uniform pulse train. The characteristic of this waveform is the distinct ambiguities and clear areas in the ambiguity function. A plot of the ambiguity function (AF) of a uniform pulse train was illustrated

in Fig.2.6. It consists of a regular array of spikes in the delay-frequency shift domain. Pulse trains eliminate the uniform sidelobe pedestal, by concentrating the ambiguity volume in isolated spikes. The clear area around the central spike implies excellent resolution performance in a dense scattering medium at the expense of severe ambiguities at delays with spacing equal to the repetition period and at Doppler shifts with spacing equal to the repetition frequency.

The big asset of pulse trains is the great design flexibility they offer. By appropriate coding of a pulse train, practically any ambiguity function can be attained. For example, as it will be shown in this chapter, by optimal frequency coding, a complete smearing of all spikes is possible that yields the thumbtack ambiguity function. There are primarily three coding techniques for manipulating the spike sidelobes of the uniform pulse train:

i) Frequency coding from pulse to pulse, referred to as FSK modulation (frequency-shift keying) or frequency hopping. This is a powerful modulation technique with a great flexibility, which allows spreading of the ambiguity surface away from the delay axis. Frequency coding will be discussed in detail subsequently.

ii) Phase coding from pulse to pulse. This modulation is referred to as PSK (phase-shift keying). Both FSK and PSK pulse trains are widely used in mobile communications. If each pulse of the train consists of more than one cycles, internal phase coding on each pulse is possible. However, this requires rather long trains and their potential in ultrasound imaging is limited.

iii) Varying the pulse repetition period, a technique known as *staggering* . Generally staggering prevents the translated ambiguity functions of each pulse from coinciding, causing effectively a smearing of the spikes.

10.2 Ambiguity function of pulse trains

Let a pulse train consist of N pulses that are separated in time by a constant pulse repetition period T_r. The complex envelope of the pulse train will then be:

$$\mu(t) = \sum_{n=0}^{N-1} \mu_n(t - nT_r), \tag{10.1}$$

where μ_n is the complex envelope of the n-th pulse normalized with respect to its energy. Let the duration of each pulse be T, with $T = T_r$ for contiguous pulses and $T < T_r$ when there are gaps between pulses. In the general case in which the train is both frequency- and phase-modulated, μ_n will be given by:

$$\mu_n(t) = \left\{ \begin{array}{ccc} \sqrt{\frac{2}{T}} e^{j2\pi f_n + j\varphi_n} & , & 0 \le t \le T \le T_r \\ 0 & , & elsewhere \end{array} \right\} \tag{10.2}$$

Using the definition of the ambiguity function from (2.28), the ambiguity function of the pulse train in (10.1) is:

$$\chi(\tau, f_d) = \frac{1}{N} \int_{-\infty}^{\infty} \sum_{n=0}^{N-1} \mu_n(t - nT_r) \times \sum_{m=0}^{N-1} \mu_m^*(t - \tau - mT_r) e^{j2\pi f_d t} dt. \tag{10.3}$$

In accordance with the definition of the ambiguity function, let the *cross-ambiguity function* between the n-th and the m-th pulse be defined as:

$$\chi_{nm}(\tau, f_d) = \int_{-\infty}^{\infty} \mu_n(t) \cdot \mu_m^*(t - \tau) \cdot e^{j2\pi f_d t} dt \ . \tag{10.4}$$

Changing the order of integration and summation and grouping summations into two categories, one obtains [68]:

$$\chi(\tau, f_d) = \frac{1}{N} \sum_{n=0}^{N-1} e^{j2\pi n f_d T_r} \times \left[\chi_{nn}(\tau, f_d) + \sum_{\substack{m=0 \\ m \neq n}}^{N-1} \chi_{nm}(\tau - (n-m)T_r, f_d) \right] \tag{10.5}$$

The latter equation shows that the ambiguity function of a pulse train is a superposition of the cross-ambiguity functions between the individual pulses, translated to the positions $\tau = (n-m)T_r$ on the delay axis and weighted by a phase factor. The desired central spike is only a superposition of the auto-ambiguity functions of all pulses. When one moves along the delay axis, the number of cross-ambiguity functions that superimpose decreases linearly with T_r, i.e. at $\tau = pT_r$ there are $N - p$ combined functions.

All the cross- and auto-ambiguity terms in (10.5) are zero for $|\tau| \geq T$. However the tails of these surfaces along the delay axis will interfere when the pulses are contiguous. The necessary condition for no interference is obviously that the duty cycle of the pulse train should be larger than 50%. If each pulse itself is a high time-bandwidth product waveform, the time extent of the cross-ambiguity term will be less than the pulse duration and lower duty cycles can be used with no interference.

10.3 FSK modulation and Costas arrays

Let each pulse in the train be a single-carrier pulse at a different frequency,i.e.

$$\mu_n(t) = e^{j2\pi f_n}. \tag{10.6}$$

The cross-ambiguity function of two contiguous pulses can be derived analytically using (10.4)[68]:

$$\chi_{nm}(\tau, f_d) = \frac{(T_r - |\tau|)}{T_r} \frac{\sin[\pi a(T_r - |\tau|)]}{\pi a(T_r - |\tau|)} \exp\left[-j\pi a\,(T_r + \tau) - j2\pi f_m \tau\right] \quad , \quad |\tau| \leq T_r \ , \tag{10.7}$$

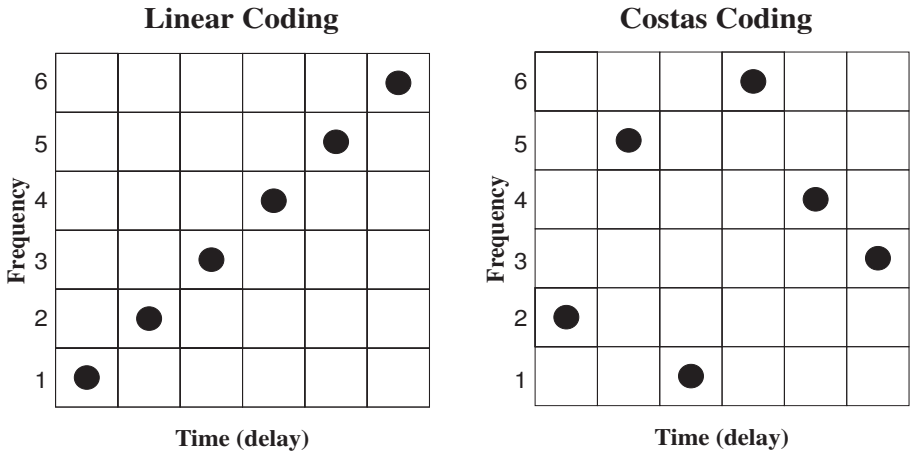

Figure 10.1: Time-frequency matrices showing the frequency firing order of linear and Costas FSK signals.

where $a = f_n - f_m - f_d$. This equation along with (10.5) show that the cross-ambiguity terms of a frequency-hopped train are translated both in delay and frequency shift away from the origin, centered at $\tau = (n-m)T_r$, $f_d = f_m - f_n$. Therefore, frequency coding of trains allows displacement of the cross-terms away from the delay axis, a property that FSK signals have but not PSK signals. The exact placement of the cross-terms depends on the order of the frequencies of each pulse.

The frequencies of each pulse can be chosen so that

$$f_n = \frac{\theta_n}{T} \tag{10.8}$$

where $\theta_n = 1,2,...N$. If all pulses have the same duration T, the integer numbers θ_n represent the number of cycles in each pulse. The pulse train will consist of pulses having N adjacent frequencies spaced $1/T$ apart, where no frequency is transmitted twice. Fig.10.1 shows the time-frequency matrix of two frequency-hopped pulse trains consisting of six pulses [69]. In the first train the frequency is stepped linearly from pulse to pulse, whereas the frequency hopping of the second train is based on the sequence $\{2,5,1,6,4,3\}$. In fact this frequency firing order is not random, but is based on a sequence known as Costas matrix. As (10.5) and (10.7) indicate, the order of frequency hopping affects the ambiguity function. A qualitative picture of the shape of the ambiguity function can be obtained by overlaying the matrices shown in Fig.10.1 with their shifted versions and counting the number of coincidences [69]. The overlay of each matrix with itself (without any shifting) has 6 hits for both cases and shows the mainlobe peak of the ambiguity function.

Linear frequency stepping from pulse-to-pulse shears the ambiguity surfaces, in a similar manner as with the linear FM signal, affecting all ambiguity spikes. For example, if the first matrix of Fig.10.1 is shifted one position right and one up, the overlay with the original matrix will give 5 hits. Two shifts right and two up will result in 4 coincidences and so on.

Similar shift of the Costas matrix will not yield more than one hit for all shifts. In fact this is the design criterion for a Costas sequence. The problem of firing order can be stated as the placement of N ones in a $N \times N$ matrix, such that each row as well as each column contains a single one. The placement should be such that for all possible x-y shift combinations of the resulting permutation matrix relative to itself, at most one pair of ones coincides. This is equivalent with the problem of ordering a set of integer numbers (corresponding to frequency indices), so that the differences between adjacent integers have no repeated values for all frequency steps. The mathematical solution of this problem is not trivial. This class of frequency hop signals were first introduced by Costas [68], who also proposed algebraic solutions to the problem. For systematic construction algorithms of Costas arrays, the reader is referred to [68] and [70].

Recall from Chapter 2 that the ambiguity surface of a single-carrier pulse is a *sinc* function along the frequency shift axis. Therefore, each cross-ambiguity surface will have a $\sin x/x$ shape whose argument is a function of the frequency differences. Along any delay line $\tau = pT_r$, the delay argument in (10.7) will be zero, and (10.7) gives the cross-term centered at $(pT_r, f_n - f_m)$:

$$\chi_{nm}(0, f_d) = \frac{\sin[\pi(f_n - f_m - f_d)T_r]}{\pi(f_n - f_m - f_d)T_r} \quad . \tag{10.9}$$

The sidelobes of all *sinc* functions for a given value of p will interfere in a complex manner because of their different phases. This is especially true for small values of p (close to the origin), where the number of cross-terms is large. Because of the overlapping sidelobes, there will always be a pedestal along the delay axis, that when the sidelobes interfere constructively will exceed $1/N$. In practical terms, the maximum sidelobe peak (worst-case) on the entire $\tau - f_d$ plane is usually around $2.1/N$, where the highest sidelobe of the auto-correlation function (delay axis) is around $1.3/N$. Weighting of the transmitted pulses can partially reduce the overlapping sidelobes [29].

Costas arrays give the best known approximation of the thumbtack AF among all frequency-coded pulse trains (FSK), since they smear the cross-ambiguity terms optimally. The sidelobe pedestal is roughly constant and the maximum sidelobe will be approximately $20\log 10(2.1/N)$. The approximate time-bandwidth product of Costas trains is N^2 [68].

Fig.10.2 shows the ambiguity function of a Costas train that can be used in ultrasound imaging. Each pulse has a duration of 3.875 μs and the train has a 50% duty cycle with a total duration of 0.228625 ms. It covers a frequency bandwidth of about 7.5 MHz and is centered around 4 MHz. It is based on a "Costas-Welch-30" array with a firing order [68]:

3,9,27,19,26,16,17,20,29,25,13,8,24,10,30,
28,22,4,12,5,15,14,11,2,6,18,23,7,21,1

The central part of the auto-correlation function of this train is shown in Fig.10.3. The mainlobe is the superposition of all the ambiguity functions of the individual pulses. The fine structure is determined by the sum of the terms $\exp[j2\pi n(f_r\tau + T_r f_d)]$. Only along the delay axis, the order of the frequencies does not play any role. Costas [68] derived the equation for the central spike along the delay axis for the case of pulses with rectangular envelope and frequency spacing equal to the

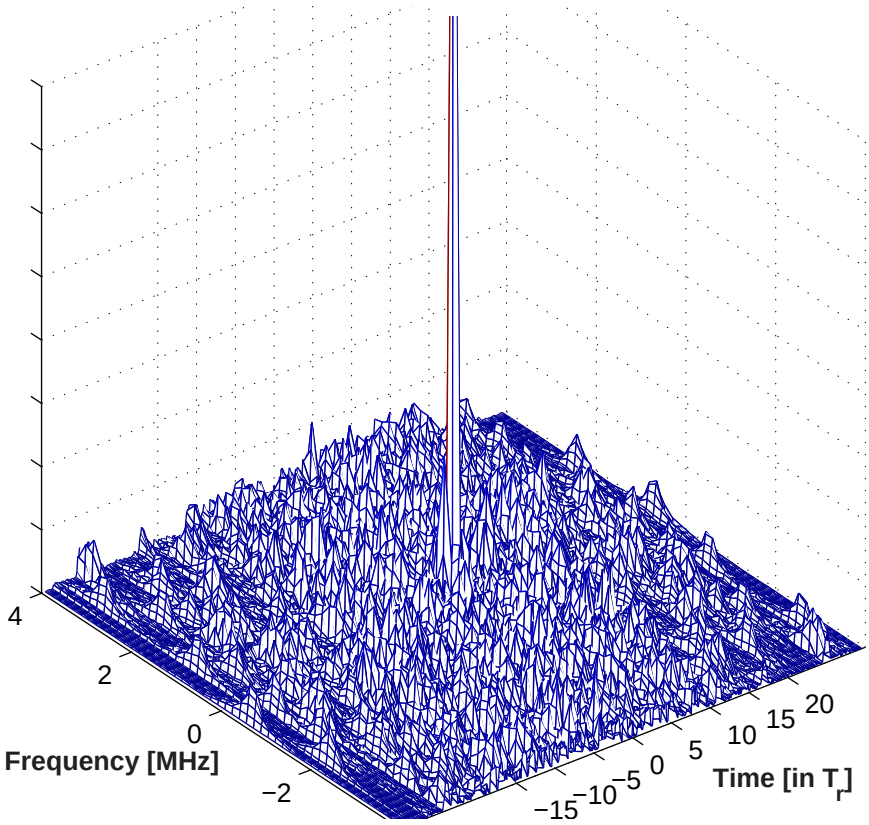

Figure 10.2: The ambiguity function of a Costas train of length 30. The mainlobe has been truncated to 70% of its maximum to reveal details of the pedestal. The peak sidelobe value of each cross-term is 20 log(1/30)=-29.5 dB below the mainlobe peak.

reciprocal of pulse duration:

$$\chi(\tau,0) = \frac{(T_r-|\tau|)}{T_r} e^{-j\pi(N-1)(\tau/T_r)} \frac{\sin(\pi N\tau/T_r)}{\sin(\pi\tau/T_r)} \tag{10.10}$$

It can be seen that this choice of pulse duration and frequency spacing is an important design parameter for obtaining a single central spike with *sinc* sidelobes. Weighting of the individual pulses can further shape the sidelobes of the central spike. For $\tau > T_r$ the sidelobes are associated only with the cross-ambiguity functions. In this region, the highest sidelobe is around -27.5 dB. If there is sufficient bandwidth, higher length Costas trains can yield even lower sidelobes. For example, a Costas-128 train would give a sidelobe level of about -40 dB below the mainlobe. Additional phase coding on the Costas trains (FSK/PSK modulation) can only slightly improve the sidelobe level [71], since that is already optimally spread. A phase code can be repeated for each frequency pulse or a length N phase code can be applied on the entire train. However, internal phase coding (several phase slots per frequency subpulse) requires longer trains.

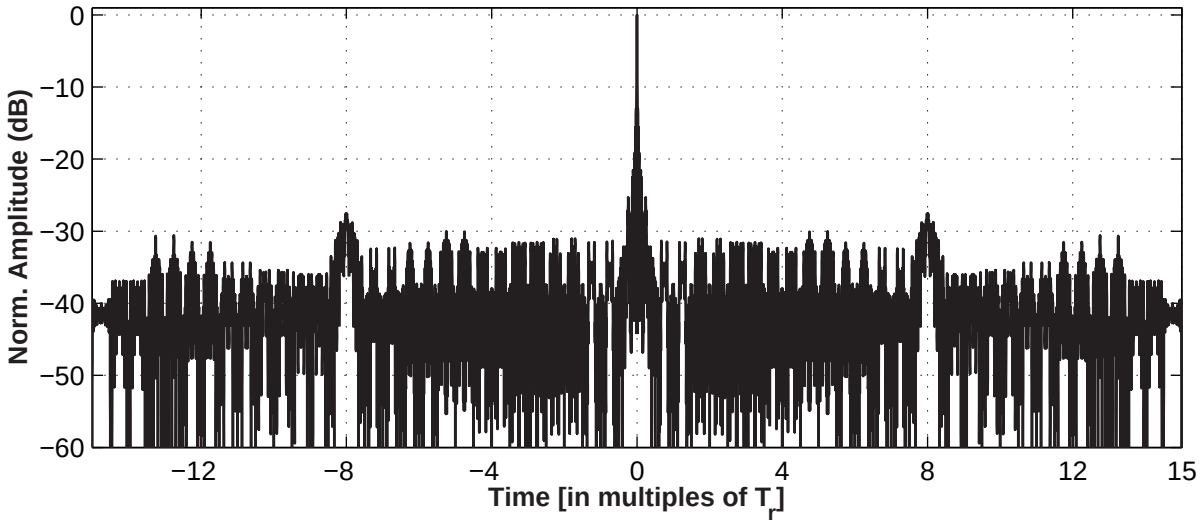

Figure 10.3: The central part of the auto-correlation function of a Costas FSK signal.

A possibility of reducing the sidelobe pedestal further is by using different Costas arrays for successive images and apply matched filtering based on a series of firings. In order for the different trains from a family of Costas trains to have good cross-ambiguity properties, the number of cross-hits for every possible delay-Doppler shift should be as low as possible. Although the number of Costas trains of a given length is high compared with other sequences, the maximum number of hits in the cross-hit arrays of pairs of Costas array depends on the number of arrays. Unfortunately, it has been shown that for the limiting case of only two arrays, the best possible case is two hits in their cross-hit array [69].

10.4 The linear FM pulse train (QLFM-FSK)

Frequency hopping can be linear from pulse to pulse. This train will be referred to as signal with QLFM-FSK (Quantized Linear FM - Frequency Shift Keying) modulation. All the cross-terms on a cut parallel to the frequency shift axis at $t = pT_r$ will now superimpose at (pT_r, pf_r), where f_r is the frequency step from pulse to pulse as well as the central frequency of the first pulse in the train. The QLFM-FSK signal translates all the sidelobes away from the delay axis by pf_r. In ultrasound, the available bandwidth does not allow a frequency step large enough to shift all p surfaces off the τ-axis and the first p surfaces will introduce some sidelobes in the measurement. The advantage is that the delay axis will be free from any pedestal and the only sidelobes will be a small number of spikes from the first p surfaces. Fig.10.4 shows a contour plot of the ambiguity function of the QLFM-FSK signal. The train consists of 32 pulses with a duty cycle of 50% and a total duration of 204 μs. Each pulse is weighted in time with a Chebyshev window. The frequency spacing from pulse to pulse is 258 KHz, covering a bandwidth of about 7.5 MHz centered around 4 MHz.

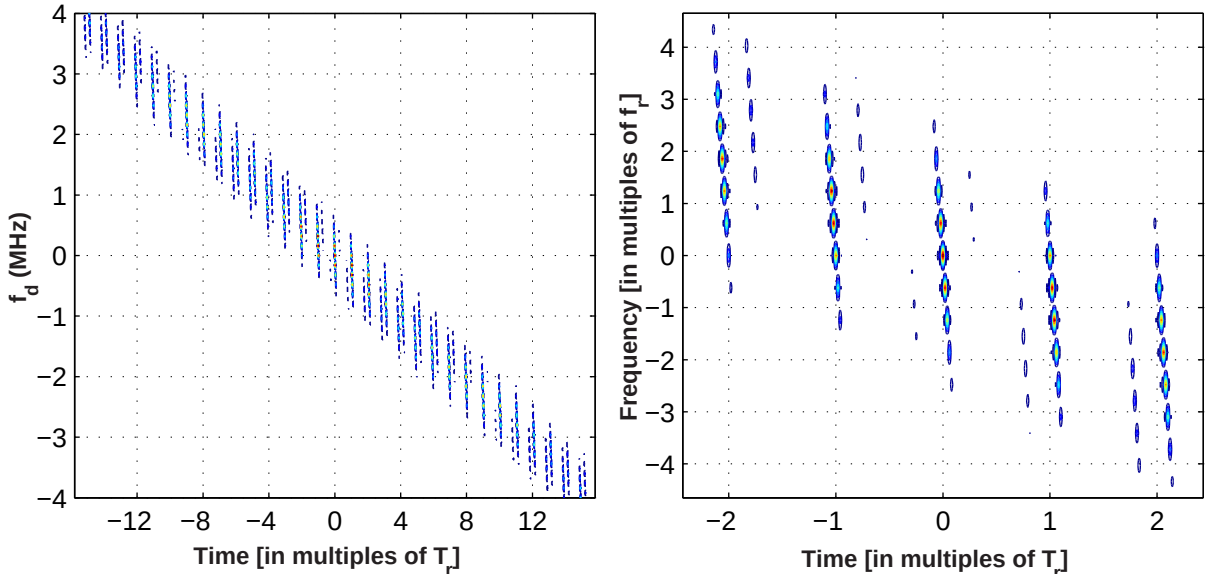

Figure 10.4: Contour plot and detail of the central part of the ambiguity function of the QLFM-FSK signal with a 50% duty cycle.

Not only the gross structure of the ambiguity function is sheared to a ridge, but also the fine structure of each p surface. The shape of the central spike is given by (10.10) but it is sheared from the delay axis.

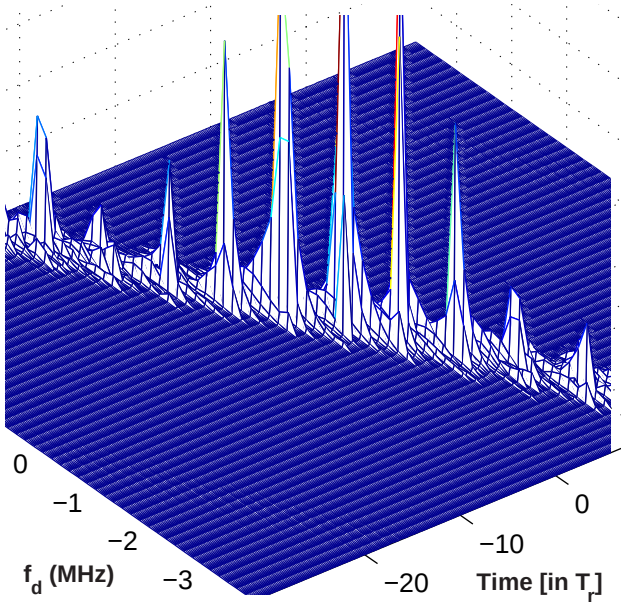

Figure 10.5: 3-D plot of the ambiguity function of the QLFM-FSK signal with a 50% duty cycle.

10.5 Fast imaging with pulse trains

One possibility for fast ultrasound imaging would be to transmit a number of Costas trains with good cross-ambiguity properties simultaneously in different directions [72]. A linear-array scan can be performed, where each group of elements is excited with a Costas signal and focused along a given line. The range sidelobes for a Costas train of length 30 would be around -30 dB, while the maximum cross-ambiguity between any two members in the train family will be -23 dB. However, similar performance can be achieved by m-sequences, which are also shorter than the pulse trains.

An alternative approach is to transmit spherical waves of a basic short pulse from every element with a constant transmit delay from channel to channel. If the inter-channel delays are long enough, the waves will not interact. At a given imaging point, all pulses will add up to a pulse train. The delays between the pulses of the received echoes will have slightly changed because of acoustic propagation. Thus, each transducer element receives a staggered pulse train. This method will be discussed in details in the following section.

10.6 A New Coding Concept

The new imaging method is based on a combined time-space coding scheme, where all elements are excited with short pulses and the high time-bandwidth product waveforms are generated acoustically.

Fig. 10.6 shows the transmitting sequence scheme. Each transducer element is excited with a

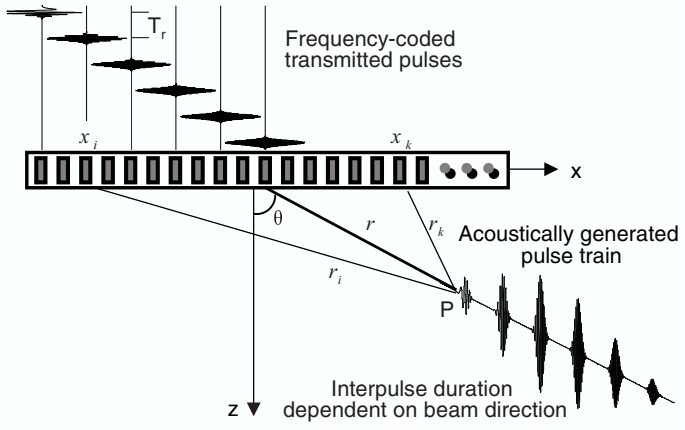

Figure 10.6: The proposed transmitting scheme results in travelling staggered pulse trains.

waveform with a constant relative delay from element to element, long enough to assure no pulse overlapping for all depths in the image. The propagating pulses will arrive at an imaging point P at different times, according to the travelling distances r_i of each pulse. Referring to Fig. 10.6, the condition that assures no pulse overlapping will be:

$$\underbrace{\sqrt{(|\vec{r}|\cos\theta)^2+(|\vec{r}|\sin\theta-x_i)^2}}_{|\vec{r}_i|}-\underbrace{\sqrt{(|\vec{r}|\cos\theta)^2+(|\vec{r}|\sin\theta-x_k)^2}}_{|\vec{r}_k|}<[(k-i-1)\,T_r+T_{ip}]\,c \tag{10.11}$$

for $k > i$, where T is the duration of each pulse, $T_{ip} = T_r - T$ is the interpulse delay, and c is the speed of sound. Expressing (10.11) as a function of the duty cycle, one obtains:

$$\text{duty cycle} \;<\; \frac{k-i}{1+\frac{|\vec{r}_i|-|\vec{r}_k|}{cT}} \tag{10.12}$$

(10.12) does not impose strict constrains on the duty cycle. For pulses 2-2.5 μs long and a transducer with $\lambda/2$ element spacing, a duty cycle of less than 97% assures no pulse overlapping of the transmitted wavefronts for all depths in the image. The duration of each transmit event in the new method is roughly N/DC higher than in conventional imaging. For N=32 transmitting pulses with a duty cycle DC=90%, the emission time is 35 times longer. The new method requires that all elements are in receive mode during the entire emission period apart from the short time in which they transmit. Neither focusing (which is an add-only operation) nor compression are affected by the short ”dead” transmitting time.

The simulation parameters for the results presented in this section are the ones listed in Table 10.1. All pulses have the same duration and thus roughly the same energy, but they have different carrier frequencies which are multiples of a basic frequency spacing parameter f_r. In this way, the first pulse consists of 1 cycle, the second has 2 cycles and so on. All pulses are weighted with a Chebyshev window. Fig.10.7 shows the frequency spectra of all transmitted pulses.

Fig.10.8 shows the echoes received from individual elements from a point scatterer located off axis at position (35.2 0 35.5) mm. The upper graph shows the excitation pulses of all elements with

Table 10.1: Simulation parameters for fast imaging using pulse trains

SIMULATION PARAMETER	VALUE
Central frequency	4 MHz
Element pitch	$\lambda/2$
Element height	10 mm
Number of elements	64
Transmit apodization	None (rectangular)
Transmit focus	None (spherical waves)
Receive aperture size	64 elements
Receive apodization	None (rectangular)
Receive focus	Two-way dynamic

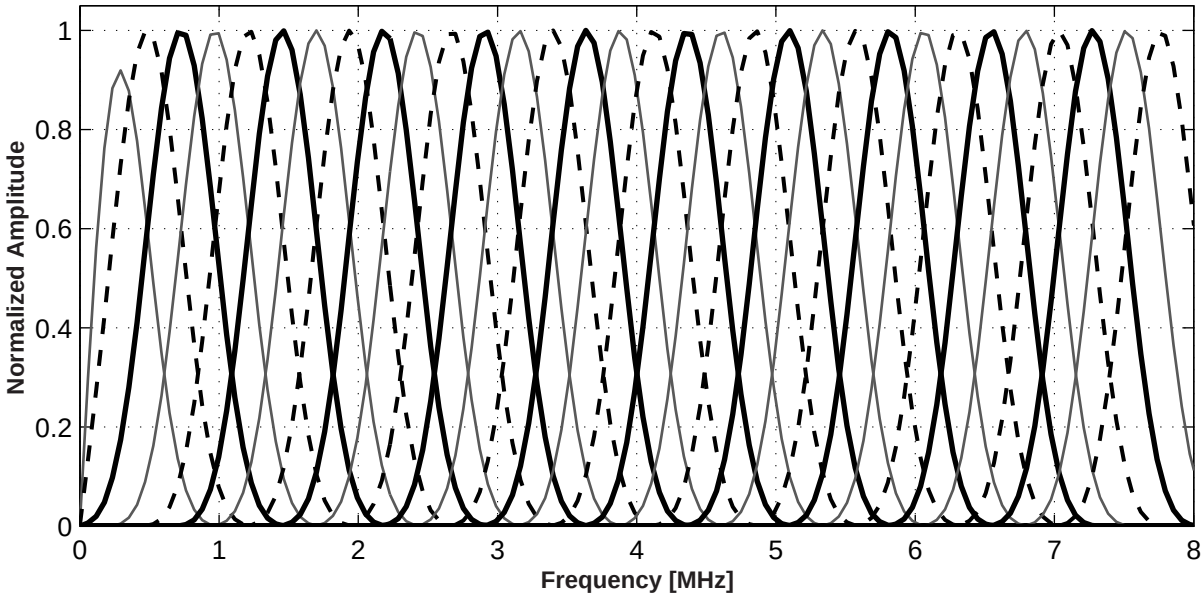

Figure 10.7: Frequency spectra of the 32 pulses transmitted from every second element of a linear array with 64 elements.

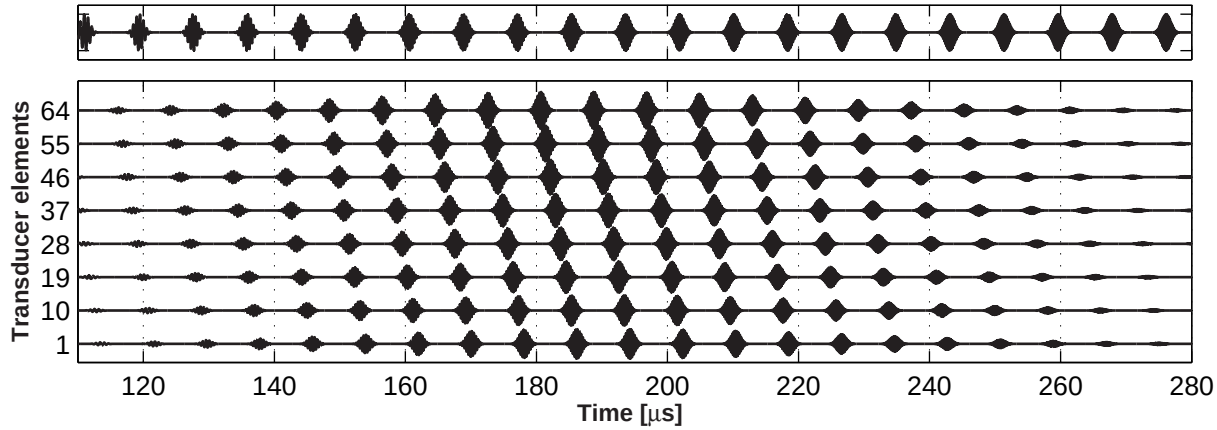

Figure 10.8: The central part of the transmitted train (when the pulses transmitted from all elements are put together) and the echoes received from individual elements from a point scatterer located at depth 5 cm 45 degrees off axis.

the applied constant delay from element to element. The interpulse delays of the received pulse train change according to the propagation delays. The interpulse periods depend on the position of the scatterer. Notice also the weighting of the received train due to the convolution of each transmitted pulse with the two-way impulse response of the transducer. The Fourier transform of a pulse train is the summation of the frequency responses of each pulse, with a frequency spacing $1/f_r$. The responses are shifted by the factor $\exp(-j2\pi(n-1)fT_r)$, and when the center frequency is a multiple of $1/T_r$, all individual responses have the same shape. For rectangular envelope pulses, the individual frequency bands are *sinc* functions. Amplitude weighting of the individual pulses allows control of the interference pattern between frequency sidelobes. Although in this case Chebyshev tapering is used for simplicity, in a practical implementation amplitude tapering should not be used because of loss in transmitting power. The shape of the frequency sidelobes can be controlled by calculating a set of phase coefficients for each pulse. This is similar to the problems of designing *moving target indication* (MTI) filters or *array pattern nulling* techniques. For design strategies, the reader is referred to the relevant literature [73, 36, 74].

Fig.10.9 shows the frequency response of the transmitted train and that of the received echoes at the first element. The particular choice of frequency spacing, pulse duration and pulse weighting results in the comb-like function of the pulse train frequency response of Fig.10.9. Staggering merely changes the position of the nulls in the spectrum. Convolution of the received pulse train with the transducer impulse response only affects the gross structure of the spectrum, applying a weight on the entire spectrum.

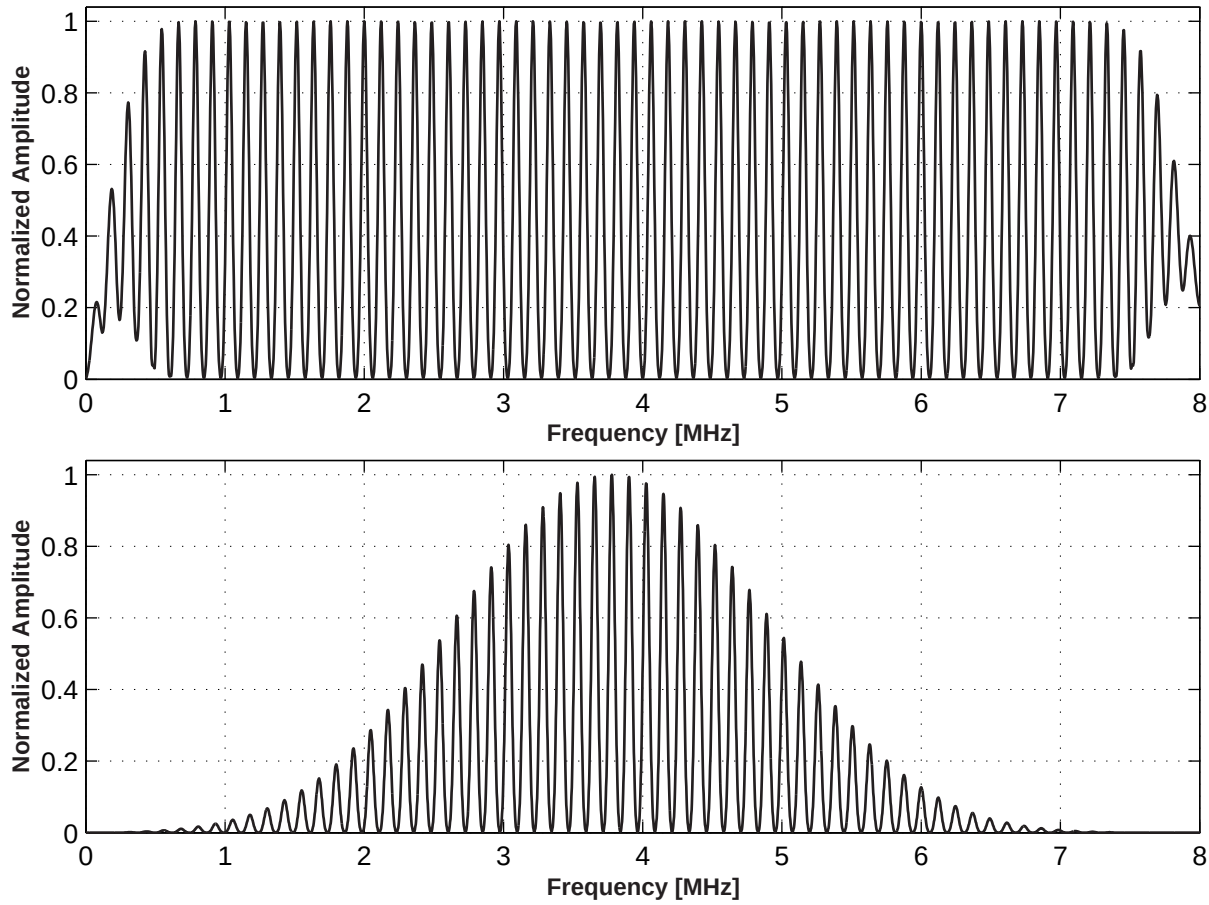

Figure 10.9: Frequency response of the transmitted train and of the received echoes at the first element.

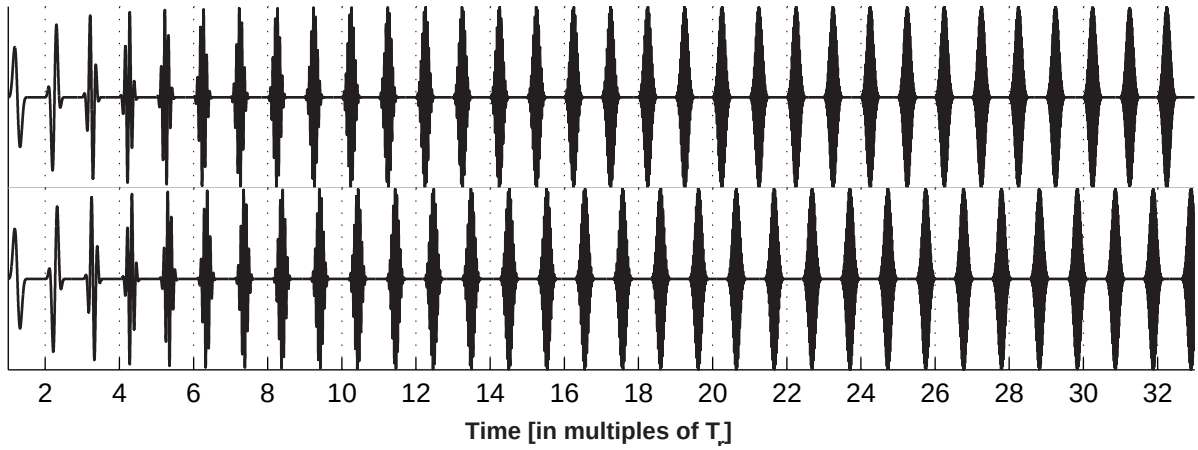

Figure 10.10: Transmitted pulse train (up) and constructed matched filter (bottom) for the beam at a 45° angle.

10.7 Coherent processing of pulse trains

The received echoes for a single point scatterer will consist of a pulse train with the number of pulses equal to the number of transducer elements. The interpulse delays are spatial-variant, consisting of the sum of the propagation delays with the applied transmit inter-element delays. However, all elements will receive a pulse train with the same set of delays from pulse to pulse. In this way, the same compression filter can be used for all channels. However a different compression filter has to be applied for every line in the image. The delays can be calculated geometrically. Ideally, a different filter should be constructed for every depth. However, compression of the QLFM-FSK is rather insensitive to staggering, and one filter for every beam direction centered at half the depth of the image is sufficient. For constructing each filter, the geometric delays are calculated, they are then translated into number of samples which are added to the constant transmit inter-pulse delays. The entire filter is a staggered pulse train, which is subsequently time reversed. For a high coarse delay precision, the pulses are resampled at 10 times the sampling frequency, the delays are calculated at the high sampling frequency and the resulting filter is downsampled in the end. Fig.10.10 shows the original transmitted train and the constructed filter for the beam at a 45° angle.

Decoding of the train is, therefore, applied on a line instead of a channel basis and prior to beamforming. The received echoes are cross-correlated with a bank of matched filters (one for each beam direction) and subsequently are beamformed for all beam directions. A new compressed rf data set is generated for every line, which is beamformed in order to generate the specific line. Matched filtering on every channel compresses the pulse train to a single pulse at the scatterer position with a number of spike axial sidelobes.

The received signal of one channel after compression is shown in logarithmic scale in Fig.10.11. This is the response from a single scatterer and the matched filter for this direction has been used for compression. By applying FSK modulation with ”per element” coding, it has been therefore

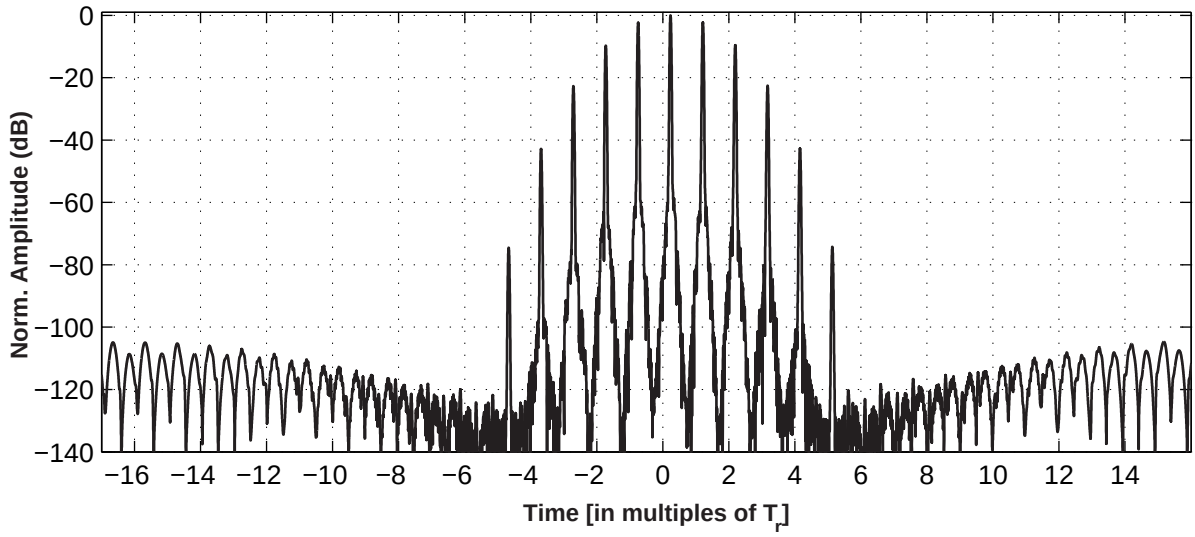

Figure 10.11: Compressed received echoes from the first element after matched filtering. This corresponds to the auto-correlation function of the staggered QLFM-FSK train received by this element.

possible to obtain rf data, which are the auto-correlation function of staggered QLFM-FSK trains.

The QLFM-FSK train described in a previous section concentrates the undesired ambiguity volume away from the delay axis apart from 4 spikes. Since the number of those spikes depends only on the frequency spacing, elimination of these spikes using a single train is not possible for realistic ultrasound bandwidth requirements. There are three basic advantages of this signal:

i) The pedestal has been entirely eliminated. From the discussion in Chapter 2, this is an important property of the ambiguity function in a dense medium.

ii) The presence of the ridge makes compression insensitive to frequency shifts. This is a general property of linear FM signals. When the received signal is frequency shifted due to attenuation, the matched filter output will still yield satisfactory compression.

iii) Staggering only shifts the sidelobe spikes.

iv) The ambiguity function between the ambiguous spikes consists of deep valleys. This can be used advantageously by combining two auto-correlation functions where the valleys of one fall in the spikes of the other, if possible.

10.8 Simulated images using pulse train excitation

Non-coherent processing

Cancellation of the spike sidelobes shown in Fig.10.11 can be done by nulling techniques, where a second set of waveforms is transmitted, that yields pulse trains with zeros at the location of the spike sidelobes. This can be obtained either by changing the frequency spacing or by changing the duty cycle of the second train. The second method has the advantage that the same pulses are transmitted from every element from one emission to the next, and therefore the effect of attenuation is the same in both images. Fig.10.12 shows the beamformed central line for two images of 3 point scatterers. For a pulse train of a given duty cycle, the duty cycle of the second train is found

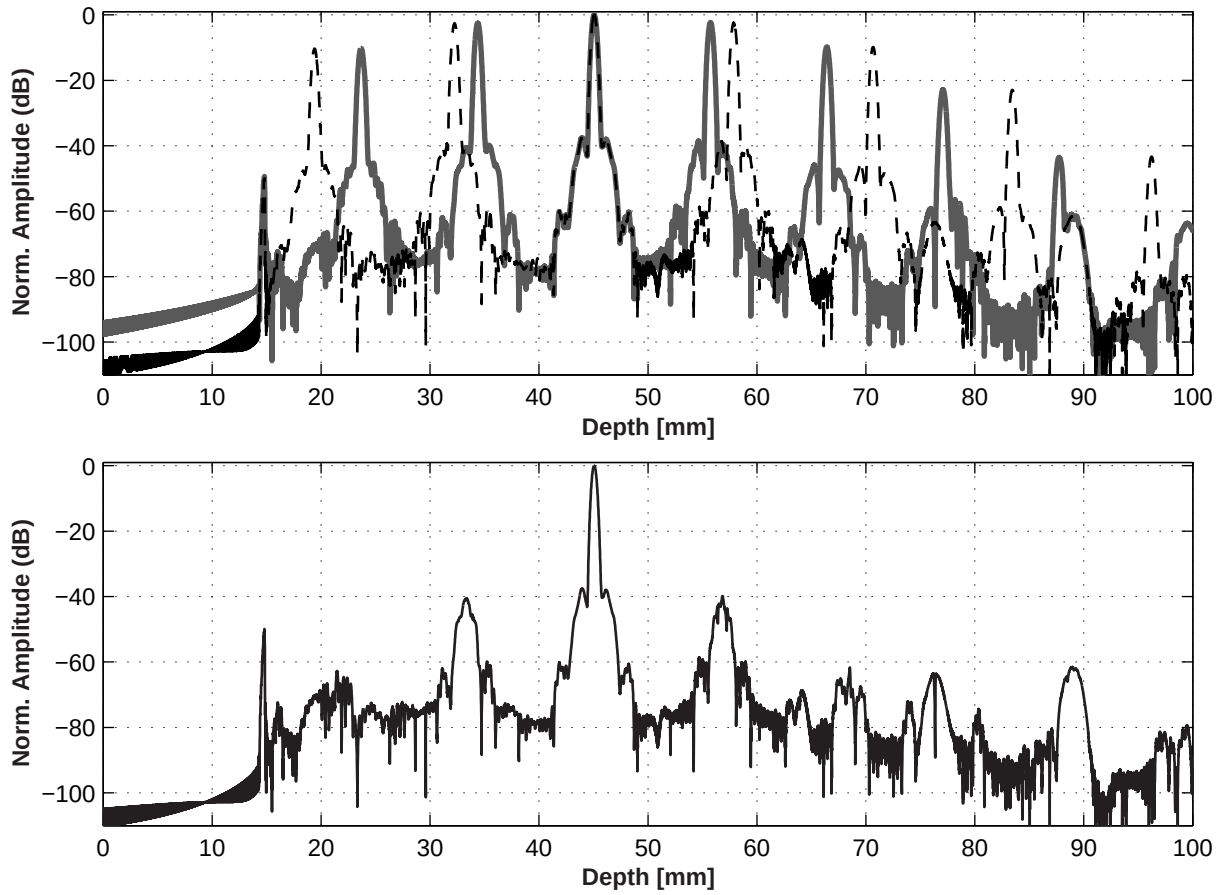

Figure 10.12: Beamformed central line for 2 different pulse train emissions with duty cycles 50% (dash black line) and 60% (solid gray line). The second graph shows the minimum of the envelope detected data from the two images.

by simply minimizing the distances for the first 5 sidelobe spikes in the auto-correlation functions of both trains. This simple technique gives a pulse train with a duty cycle of 60% when the first pulse train has a duty cycle of 50%. Cancellation of the spikes is then achieved by taking the

minimum of the envelope-detected data for every pixel in the image. In this way, high resolution images are possible with coherent combination of two images obtained by only two transmissions. In Fig.10.13 simulated single-emission images are shown with pulse train excitation. The same QLFM-FSK pulse trains are used for the first two images, with the only difference the interpulse period. The third image is the non-coherent combination of the two first images, using the min operator described above.

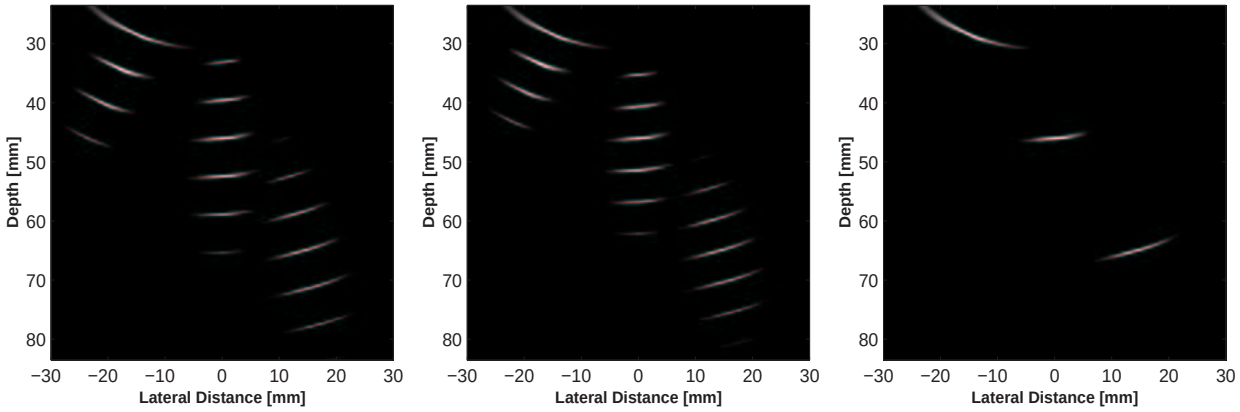

Figure 10.13: Simulated images of a phantom consisting of 3 point scatterers. The two first images are the single-emission images from pulse train excitation. In the first image the transmitted pulse train has a duty cycle of 50% and in the second image the transmitted train has a duty cycle of 60%. Because of the different duty cycles, the ambiguous spikes are in different positions and can be eliminated by taking the min of the envelope-detected data from the two images (shown in the third image). The dynamic range of all 3 images is 40 dB.

Coherent processing

Alternatively, cancellation of the ambiguity spikes would be possible for two responses which have the same mainlobe, and sidelobes equal in amplitude but with opposite phase. This requires that the second transmitting train is a QLFM-FSK signal with the same frequency coding as the first one, with additional phase coding from pulse to pulse (such trains are referred to as *FSK/PSK* signals). According to (10.5), the build up of the first spike, for instance, is due to the superposition of the cross-terms given by (10.7) of all pulses with $n - m$=1, i.e. for adjacent pulses. Setting every second pulse in the train to be 90° out of phase will add a minus sign to all cross-terms of (10.7). With this simple phase coding all pulses with $n - m$=3 will also be out of phase, which will yield cross-terms with minus sign that contribute to the third ambiguity spike. Thus, all odd-numbered spikes will have the same response, but inverted sign compared to the response of the first train. This is shown in Fig. 10.14. The plots show the beamformed and compressed rf-line at an angle of 45° for two pulse train emissions with a duty cycle of 94%. The medium consists of one scatterer at a depth of 8 cm along that beam direction.

Since phase coding eliminates only the odd-numbered spikes, the trains have to be designed in

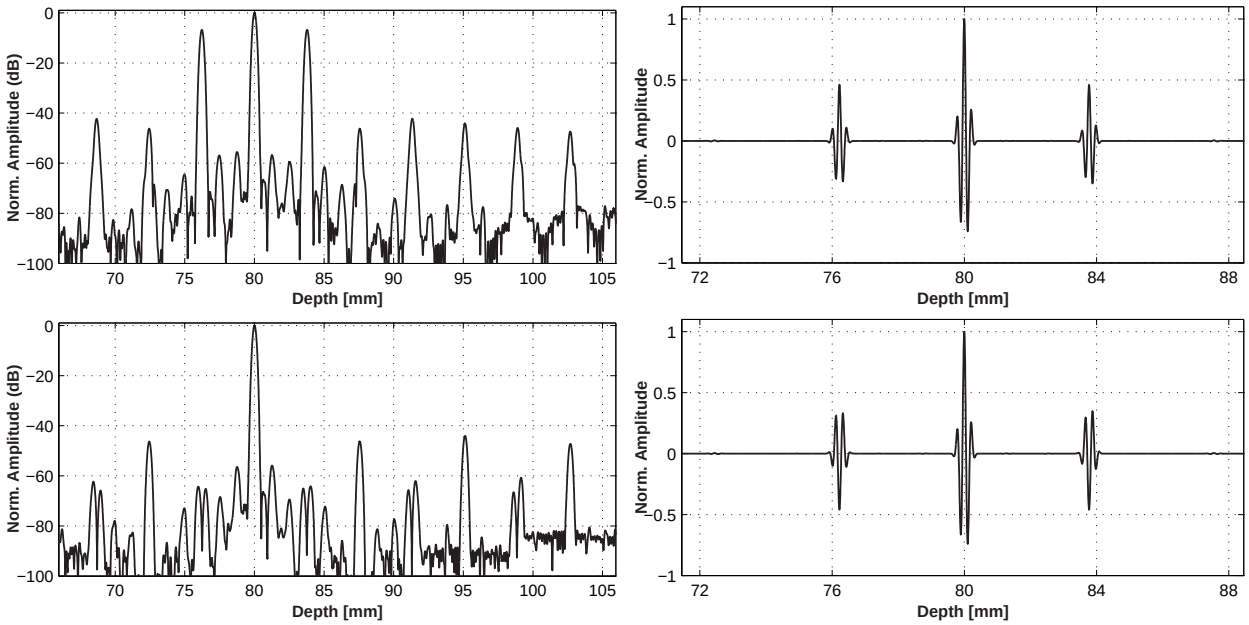

Figure 10.14: Beamformed and compressed rf-line for a QLFM pulse train emission shown in logarithmic scale at the upper left graph. A second train with additional PSK modulation will give rf-data with the same main response but ambiguity spikes with opposite phase (right graphs). Coherent sum of the rf-data from the two emissions will cancel all the odd-numbered ambiguity spikes (bottom left graph).

such a way that no even-numbered spike enters the delay axis for a desired dynamic range. The upper left graph of Fig. 10.14 and Fig. 10.11 show two different designs. In the one in Fig. 10.11, there are 4 spikes over 45 dB on each side of the auto-correlation function, while the rest of the spikes are entirely eliminated. On the other hand, the auto-correlation function of Fig. 10.14 has only 1 spike (that can be eliminated by the phase coding described before) with the price to pay that all other spikes are around 45 dB. The design parameters are the pulse weighting functions, the frequency spacing and the duty cycle of the train. Equation (10.7) shows that all cross-terms for rectangular pulses are *sinc* functions weighted by a phase factor. Their superposition can reinforce or cancel the sidelobes of the individual *sinc* functions. Weighting of the pulses can position the nulls at a given p position. Full cancellation of the sidelobes at all p positions is not possible, giving the trade off of the design.

Simulated images using the proposed method are shown in Fig. 10.15. The first image uses pulses with FSK modulation. In the second emission, all transmitted pulses are identical to the ones used in the first emission, except that every second pulse is phase-inverted. This approach has the advantage that the same pulses are transmitted from every element from one emission to the next, and therefore the effect of attenuation is the same in both images.

Cancellation of the ambiguity spikes down to -45 dB is then achieved by coherently summing the processed rf-data for every pixel in the image. This is shown in the third image of Fig. 10.15. In this way, a ghost-free image is derived by coherent summation of two images obtained by

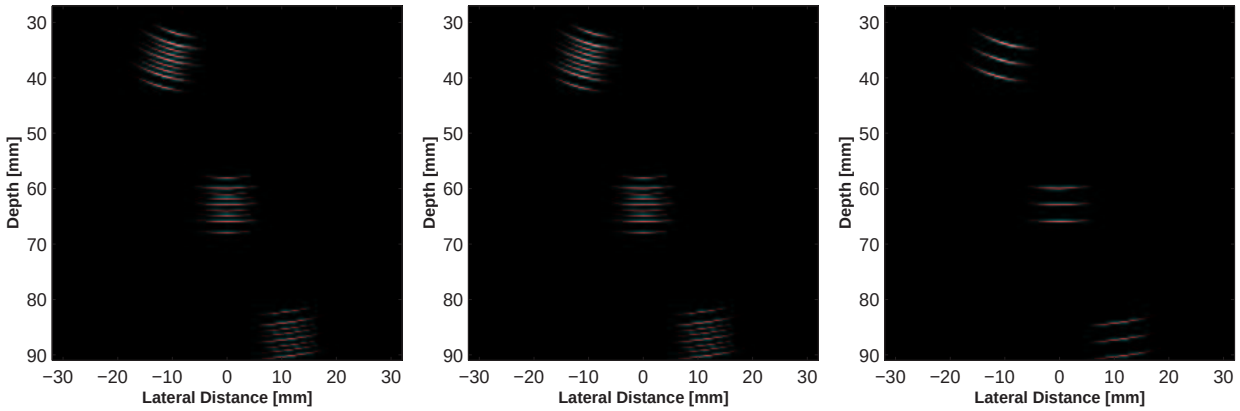

Figure 10.15: Simulated images of a phantom consisting of 3 groups of 3 point scatterers each, 3 mm apart. The two first images are the single-emission images from pulse train excitation. Both transmitted pulse trains are QLFM-FSK signals with a duty cycle of 94%. In the second image the transmitted pulse train has additional PSK modulation. Because of this difference, the ambiguous spikes are opposite in phase and can be eliminated by summing the rf-data from the two images (shown in the third image). The dynamic range of all 3 images is 45 dB.

two successive transmit events. In an actual implementation, the rf-data (after beamforming and compression) from emission $n-1$ are stored in memory, and are added to the rf-data of the current emission n. The new ghost-free image is displayed, and the rf-data of emission n are pushed to the RAM. Using this recursive scheme as first suggested by Nikolov et.al.[75], a new image can be formed after every emission.

10.9 Possible alternative imaging strategies

The interpulse periods of the train will be space-variant, which is equivalent to staggering. Therefore, the coding of the transmitted pulse train has to be as insensitive to staggering as possible. Costas arrays have already the optimal spreading of cross-ambiguity functions for all pulses and staggering can only cause some overlaps. Additionally, the constant pedestal of the Costas train is still high for the axial resolution requirements of ultrasound imaging. Then the QLFM-FSK signal is possibly the best we can do: eliminate the whole pedestal except 4 spikes on the delay axis and then resolve these 4 ambiguities.

Possible phase coding from pulse to pulse can be an additional design parameter (QLFM-FSK/PSK signals). Such coding can give the possibility of controlling the position of the ambiguous spikes without smearing them. Another idea would be to use Golay-coded pulse trains. However, the complementarity property of these trains is based on the exact position of the sidelobes and staggering of Golay-PSK trains can severely distort it.

In conclusion, the simulation results presented in this chapter serve as a proof of concept of pulse

train excitation with per element coding. However several problems associated with the new technique have to be examined. For instance, alternating between linear increasing and decreasing frequencies from emission to emission should be investigated, as well as the effect of attenuation on train decoding. Pulse trains can offer SNR improvement of 10 to 20 dB, since the time-bandwidth product of the frequency-coded pulse train is approximately equal to N^2, where N is the number of frequency bins. Pulse trains can also potentially increase the frame rate in ultrasound imaging. The poor lateral resolution of the images shown in Fig. 10.13 indicate that more advanced pulsing and beamforming techniques are required. Imaging schemes using the potential benefits of pulse trains are still to be investigated. In synthetic pulse train imaging, focusing delays should associate a pulse train to a virtual excited element and the whole transmit aperture should be synthesized with only one emission. Variable interpulse periods from element to element should be investigated. Alternatively, pulse trains could be used in coded linear array imaging as discussed in Chapter 9, where each sub-aperture transmits a different focused train.

CHAPTER

ELEVEN

Conclusions

A number of techniques for using coded signals in medical ultrasound have been described, which can increase the frame rate and signal-to-noise ratio of the resulting images. Inspired by techniques from radar, it has been described how FM-modulated signals can be emitted by conventional ultrasound transducers and subsequently compressed to yield the same axial resolution and contrast and at the same time increase the signal-to-noise ratio by 10 to 20 dB. The properties of the signals are theoretically derived, and it is shown how the coded signals can be evaluated using ambiguity functions. It is described how the higher demands on the codes in medical ultrasound can be met by amplitude tapering of the emitted signal and by using a mismatched filter during receive processing to keep temporal sidelobes below 60 to 100 dB. Other coded signals like Golay and binary codes have also been considered and characterized in terms of axial resolution, signal-to-noise ratio and temporal sidelobes. One of the main results is the conclusion that linear FM signals have the best and most robust features for ultrasound imaging. These results are demonstrated both through computer simulations and phantom and *in-vivo* measurements. The dissertation has also investigated the use of codes for increasing the frame rate in ultrasound imaging. It is shown that FM codes can be used to increase the frame rate by a factor of two without a degradation in image quality and by a factor of 5, if a slight decrease in image quality can be accepted. The use of synthetic aperture imaging is also considered, and it is here shown that Hadamard spatial encoding in transmit with FM emission signals can be used to increase the frame rate by 12 to 25 times with either a slight or no reduction in signal-to-noise ratio and image quality. By using these techniques a complete ultrasound phased array image can be created using only two emissions. A preliminary investigation of pulse trains is also carried out and has shown the potential for even more improvements.

APPENDIX

A

Relevant publications

A.1 Potential of coded excitation in medical ultrasound imaging

Potential of coded excitation in medical ultrasound imaging

Thanassis X. Misaridis[1], Kim Gammelmark[1], Christian H. Jørgensen[1],
Niklas Lindberg[1], Anders H. Thomsen[1], Morten H. Pedersen[2], Jørgen A. Jensen[1]

[1]Center for Fast Ultrasound Imaging,
Department of Information Technology, Build. 344,
Technical University of Denmark,
DK-2800 Lyngby, Denmark
[2]Herlev University Hospital, Department of Ultrasound, DK-2730 Herlev, Denmark

Published in Ultrasonics, Elsevier Science B.V., 2000.

POTENTIAL OF CODED EXCITATION IN MEDICAL ULTRASOUND IMAGING

Thanassis X. Misaridis[1], Kim Gammelmark[1], Christian H. Jørgensen[1],
Niklas Lindberg[1], Anders H. Thomsen[1], Morten H. Pedersen[2], Jørgen A. Jensen[1]

[1]Center for Fast Ultrasound Imaging, Department of Information Technology, Build. 344,
Technical University of Denmark, DK-2800 Lyngby, Denmark
[2]Herlev University Hospital, Department of Ultrasound, DK-2730 Herlev, Denmark

Abstract

Improvement in signal-to-noise-ratio (SNR) and/or penetration depth can be achieved in medical ultrasound by using long coded waveforms, in a similar manner as in radars or sonars. However, the time-bandwidth product (TB) improvement, and thereby SNR improvement is considerably lower in medical ultrasound, due to the lower available bandwidth. There is still space for about 20 dB improvement in the SNR that will yield a penetration depth up to 20 cm at 5 MHz [1]. The limited TB additionally yields unacceptably high range sidelobes. However, the frequency weighting from the ultrasonic transducer's bandwidth, although suboptimal can be beneficial in sidelobe reduction.

The purpose of this study is an experimental evaluation of the above considerations in a coded excitation ultrasound system. A coded excitation system based on a modified commercial scanner is presented. A predistorted FM signal is proposed in order to keep the resulting range sidelobes at acceptably low levels. The effect of the transducer is taken into account in the design of the compression filter. Intensity levels have been considered and simulations on the expected improvement in SNR are also presented. Images of a wire phantom and clinical images have been taken with the coded system. The images show a significant improvement in penetration depth while they preserve both axial resolution and contrast.

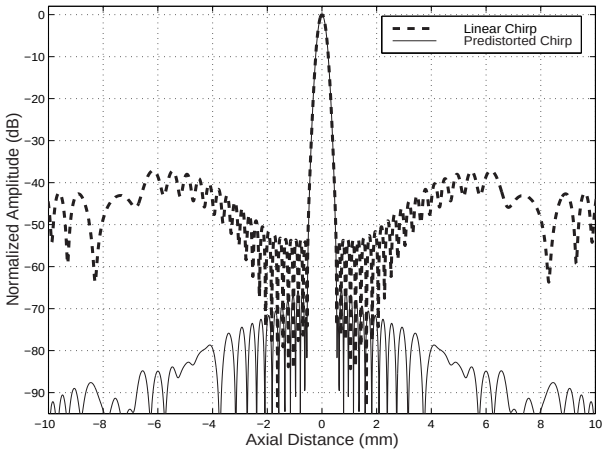

Figure 1: The effect of the transducer and predistortion on sidelobe level and energy. The figure shows the compressed output of two chirps 20 μs long. The impulse response of the transducer has been simulated as two cycles of the 4 MHz sinusoid signal weighted by a Hanning window.

1 Introduction

Pulse compression theory and techniques were developed mainly for radar systems and adjustments to the ultrasound specific problem have to be considered carefully. M-sequences and polyphase codes that have been used in radars are based on abrupt changes in phase that contain high harmonics, that the ultrasonic transducer smoothes out. Therefore, pulse compression with binary codes applied in ultrasound systems has worse performance than the theoretical expected and have not been considered further. Instead, linear chirps of 20 μs in length have been used.

Range sidelobes represent an inherent part of the pulse compression mechanism, primarily due to the abrupt rise in the signal spectrum. In radar systems, the most used approach for sidelobe reduction is frequency weighting of the output spectrum. Among a plethora of window functions,

the Dolph-Chebyshev window exhibits the lowest sidelobes and has been used extensively in radars [2]. The efficiency of frequency windowing is based on the concept of shaping a nearly rectangular spectrum to reduce the ripples of the sinc function in the time domain. While the approximation of a rectangular spectrum holds in radar systems, it is not valid in the low time-bandwidth chirps that can be used in medical ultrasound. In ultrasound, however, suboptimal windowing occurs naturally from the transducer. Thus, if we design the matched filter to take into account the transducer frequency weighting, the sidelobe levels of the compressed signal are significantly reduced. Fig. 1 shows that pulse compression of a linear chirp with a modified mismatched filter, gives an output signal with 37 dB range sidelobe level to start with. The -20 dB axial resolution achieved after compression is 0.641 mm. For comparison, the axial resolution for the conventional pulse is 0.620 mm. It can be observed that the effect of the transducer on a linear chirp does not remove the distant sidelobes. The reason for that is their different origin. The distant range sidelobes are paired-echo distortions due to deviation from the ideal rectangular spectrum amplitude and are not affected by frequency weighting. When range sidelobes below -40 dB are required (which is the case for the typical dynamic range of 50 dB or greater of an ultrasound image), methods that compensate for that have to be employed. Additionally, the presence of these distant sidelobes, increases the mainlobe-to-sidelobe energy ratio (MSR), an important image quality measure which describes the leakage of energy from bright into dark areas of the image [3]. Data correlating pulse rise-time and spectrum ripple have been previously reported [4]. Based on this analysis, a modified chirp with shaping of the transmitted signal has been generated, that effectively reduces the spectrum ripples. Details on the coded waveform design is planned to be published in a separate correspondence. Fig. 1 shows that application of the filter on the predistorted chirp achieves a big improvement in MSR and highest range sidelobes down to -65 dB. Further reduction of the sidelobe levels with an 0 loss in axial resolution can be attained if Chebyshev frequency weighting is applied subsequently. Thus, an implementation of the new scheme can easily give the flexibility to switch between finer resolution and minimum range sidelobes, depending on the application and the user's needs. The new excitation/compression scheme has been tested with Field II, a program for simulating ultrasound fields [5]. A wire phantom has been simulated and the expected performance in terms of resolution and range sidelobe levels has been verified. The new scheme is also not sensitive to the effect of frequency shift due to attenuation.

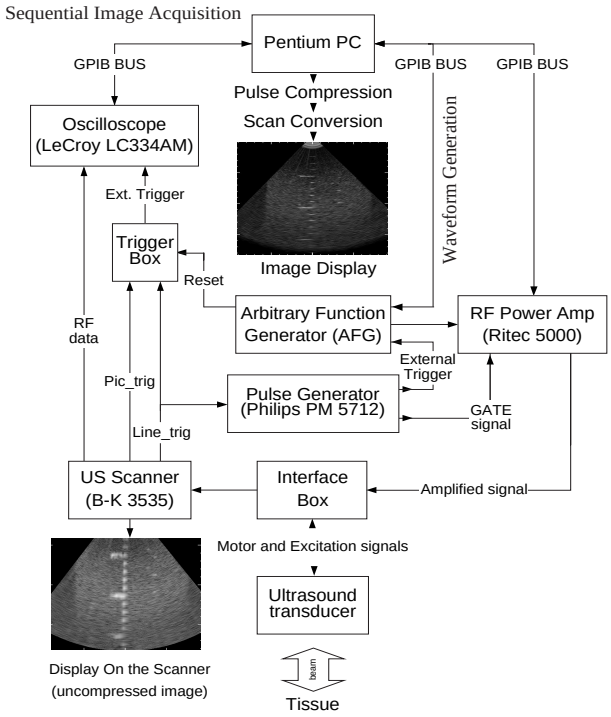

Figure 2: Block diagram of the experimental setup.

2 Materials and Methods

A block diagram of the coded excitation system is shown in Fig. 2. A commercial scanner (B-K Medical 3535) was used along with a single-element mechanically rotating 4 MHz transducer. The transmitter of the scanner was inactivated and the machine was modified to operate with external excitation. A LeCroy 9112 arbitrary function generator (AFG) was used for the generation of the coded waveforms on the computer. A power RF amplifier (Ritec 5000), specifically designed to drive ultrasound transducers, was used subsequently for amplification. The modified system was used for both conventional imaging using pulse excitation and for chirp excitation. The acquisition of the RF data was done using the sequential acquisition capabilities of an 8-bit digital oscilloscope (LeCroy LC334AM) at a minimum sampling frequency of 20 MHz depending on the requested depth of penetration and number of scan lines. The transfer time of raw binary data for an image with 105 scan lines and a depth of 18 cm was 0.22 s, allowing acquisition rate of 4.5 frames/sec. The total processing and display time was 2.714 s/image. Both the generation of the signals through the AFG and the image acquisition from

the Oscilloscope were controlled remotely from the computer through a GPIB interface. Software libraries were written for both devices, allowing high level programming through MATLAB^{TM}. After acquisition of the RF data, all post processing and display is done on the computer. That includes high-pass filtering, pulse compression, interpolation and scan conversion, logarithmic compression, envelope detection and display. The whole system was running under Linux and could be operated from any computer connected to the network.

Both clinical images and images of a wire phantom were taken for the evaluation of the proposed coded scheme. The phantom consisted of wires placed every 1 cm and attenuated medium with attenuation of 1 dB/[MHz×cm].

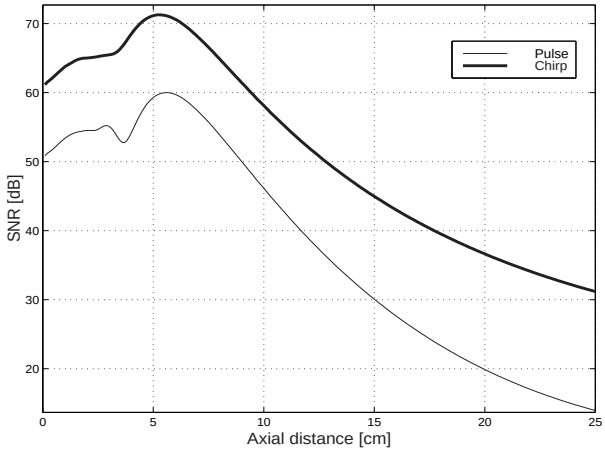

Figure 3: Simulation results on the expected SNR improvement from coded excitation using Field II. There is a 12 to 17 dB improvement depending on the depth.

3 Intensity and SNR considerations

Simulations using Field II, as well as hydrophone measurements in a water tank were performed for intensity calculations. Applying pulse and coded excitation of equal peak voltage amplitude resulted in a measured I_{sptp} (spatial-peak-temporal-peak) about 1.6 times greater for the coded signal at the acoustical focus. This is due to different focusing properties of the transducer for a much longer signal. The relationship is non-linear and is between 1.1 and 1.8 depending on the applied voltage amplitude. For instance, a conventional pulse of 80 V and a chirp signal of 50 V give the same I_{sptp} at the focal point. The excitation voltages were normalized to yield the same I_{sptp} for both sorts of excitation signals.

The FDA limits for adult cardiac imaging [6] are I_{spta}=730 mW/cm^2 (spatial-peak-temporal-average) and I_{sppa}=240 W/cm^2 (spatial-peak-pulse-average) in water. For the transducer used, the intensities measured in water by the company with a conventional pulse of 70V were I_{spta}=3.19 mW/cm^2 and I_{sppa}=66.4 W/cm^2 at the focal point. Both intensities are well below the limits. I_{sppa} is the mean intensity over the pulse duration at the focal point and therefore does not change significantly, when the duration of the signal changes. However, the I_{spta} is an integration over time and is increased when the duration of the pulse is increased. For a chirp of duration 20 μs, the I_{spta} is about 60 times higher depending on the applied code, i.e. 196 mW/cm^2, which is still 3.7 times lower than the specified safety limits.

Simulations were performed with Field II to investigate the improvement in SNR. The transducer used for the experiments was simulated (nominal frequency 3.82 MHz, 67% -6 dB bandwidth). An attenuating medium with frequency-dependent attenuation of 0.7 dB/[MHz×cm] was used. The SNR calculation has been based on the following equation:

$$SNR_i = \frac{\int_0^{T_{PRF}} p_i^2(t,\vec{r}_i)dt}{P_{noise}}, \qquad (1)$$

where p is the pressure field at the axial position $\vec{r}_i$, T_{PRF} is the period from pulse to pulse and P_{noise} is the power of the band-pass thermal noise of the system. The excitation signals have been normalized to yield the same I_{sptp} at the focal point and the energy of the received echoes has been calculated as an integration over time of the square pressure field. Fig. 3 shows the result of the simulations normalized to a maximum of 60 dB SNR for the conventional pulse. We see that there is a SNR improvement between 12 and 17 dB for all depths. For instance, at a depth of 18 cm, an improvement of 16 dB is expected. The penetration depth can, therefore, be increased by 7 to 10 cm at 4 MHz.

4 Experimental Results

Clinical images as well as images of a wire phantom have been taken using the coded excitation system. All images shown have a dynamic range of 50 dB. Fig. 4 shows clinical images of the gallbladder from a healthy volunteer. The applied voltages have been normalized to yield the same I_{sptp} at the focal point. The decrease in noise for higher depths is apparent. Some RF lines were corrupted with spike noise resulting in brighter stripes along the image. It was found

that there was a sampling error in the oscilloscope at a certain sampling frequency and this frequency was not used subsequently. The spike noise was removed by a median filter whenever it occurred.

Fig. 5 shows images of a wire phantom with attenuation of 1 dB/[MHz×cm]. The central RF line of each image is shown on the left. The peak excitation voltages are 32 V p-p for the conventional pulse and 20 V for the chirp, that yield the same I_{sptp}. The applied excitations are very low compared to what is commonly used in ultrasound scanners. Therefore, the noise level in the conventional image on the top of Fig. 5 is high. However, when a chirp excitation is applied, there is a significant reduction in noise level for depths over 10 cm. The time-gain compensation was kept low and at the same level for both images. The 20 dB axial resolution, measured at the scatterer located at about 6 cm, was 0.71 mm for the conventional pulse excitation and 1.04 mm for the chirp. In the latter, the axial resolution depends on the applied compression filter and can be improved at the expense of higher axial sidelobes.

5 Conclusion

A new coded excitation/compression scheme based on a predistorted FM signal has been presented in this paper. The scheme has been implemented and both clinical and phantom images have shown the advantages of coded excitation in terms of improvement on the SNR and the penetration depth. The penetration depth can potentially be increased by 7-10 cm using a 4 MHz transducer. The range resolution achieved is comparable to that of a conventional system. The range sidelobes, at the same time, are well below the limits of the typical dynamic range of an ultrasound image. The energy of the sidelobe region is also reduced and therefore the image contrast is improved, by lowering the distant sidelobes caused by the ripples of the spectrum's amplitude. The new scheme has shown good stability to the effect of frequency shift due to attenuation.

6 Acknowledgements

The authors would like to thank technician Kjeld Martinsen for building one of the phantoms and transducer holders, and the Ph.D. student Peter Munk for his valuable advises.

This work was supported by grant 9700883 and 9700563 from the Danish Science Foundation and a grant from B-K Medical.

References

[1] M. O'Donnell, *"Coded Excitation System for Improving the Penetration of Real-Time Phased Array Imaging Systems"*, IEEE Trans. on Ultras.,Ferroel.,Freq. Contr., Vol.39, No. 3 (1992) 341-351.

[2] M. Skolnik, *"Radar Handbook"*, McGraw-Hill,Inc., 2nd Ed. 1990.

[3] J. W. Adams, *"A New Optimal Window"*, IEEE Trans. on Signal Processing, Vol.39, No.8 (1991) 1753-1769.

[4] C. E. Cook and J. Paolillo, *"A Pulse Compression Predistortion Function for Efficient Sidelobe Reduction in a High-Power Radar"*, Proceedings of the IEEE , Vol. 52 (1964) 377-389.

[5] J. A. Jensen, *"Field: A program for simulating ultrasound systems"*, Med.Biol.Eng.Comp., 10th Nordic-Baltic Conference on Biomedical Imaging, Vol.4, Supplement 1, Part 1 (1996b) 351-353.

[6] AIUM *Safety considerations for diagnostic ultrasound. Technical report.* American Institute of Ulrasound in Medicine, Bethesda, Maryland, 1988.

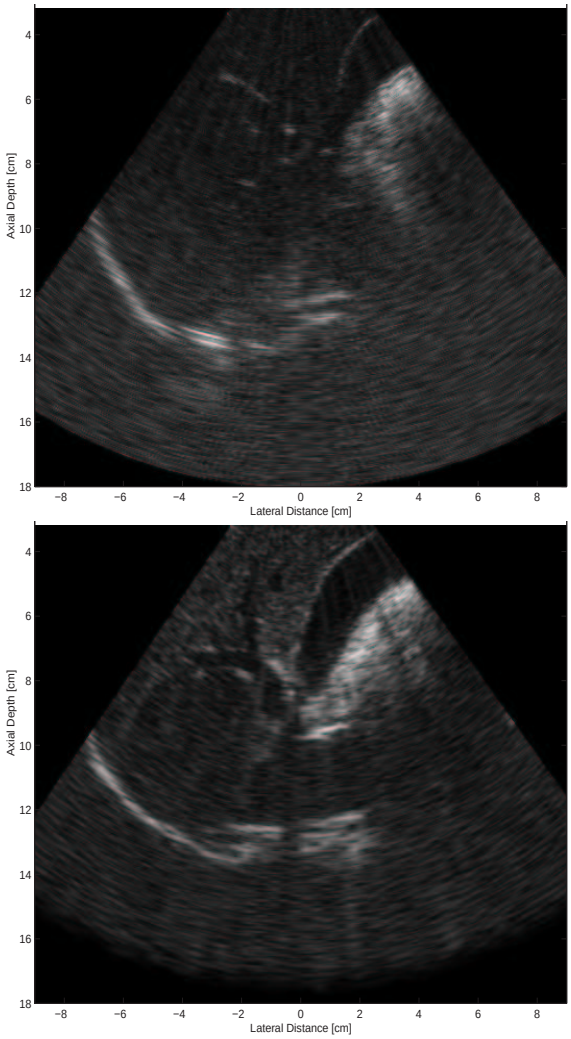

Figure 4: Clinical images of the gallbladder from a healthy volunteer that demonstrate the gain in penetration depth using codes. The applied voltages have been normalized to yield the same I_{sptp} at the focal point

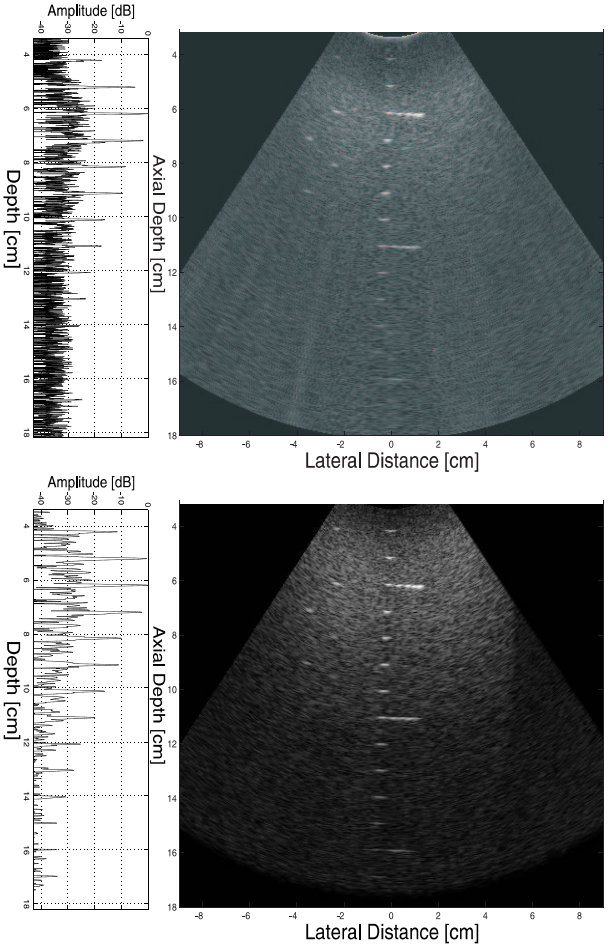

Figure 5: Images of a wire phantom with attenuation of 1 dB/[MHz×cm]. The peak excitation voltages are 32 V for the conventional pulse and 20 V for the chirp. The plots on the left side are the central RF lines of the images.

A.2 An effective coded excitation scheme based on a predistorted FM signal and an optimized digital filter

Paper presented at the IEEE International Ultrasonics Symposium, Lake Tahoe, USA, 1999.

An effective coded excitation scheme based on a predistorted FM signal and an optimized digital filter

Thanassis Misaridis and Jørgen Arendt Jensen,
Center for Fast Ultrasound Imaging,
Department of Information Technology, Build. 344,
Technical University of Denmark,
DK-2800 Lyngby, Denmark

Published in Proceedings of IEEE International Ultrasonics Symposium, Lake Tahoe, USA, 1999.

An effective coded excitation scheme based on a predistorted FM signal and an optimized digital filter

Thanassis X. Misaridis and Jørgen A. Jensen

Center for Fast Ultrasound Imaging, Department of Information Technology, Bldg. 344, Technical University of Denmark, DK-2800 Lyngby, Denmark

Abstract

This paper presents a coded excitation imaging system based on a predistorted FM excitation and a digital compression filter designed for medical ultrasonic applications, in order to preserve both axial resolution and contrast. In radars, optimal Chebyshev windows efficiently weight a nearly rectangular spectrum. For the small time-bandwidth (TB) products available in ultrasound, the rectangular spectrum approximation is not valid, which reduces the effectiveness of weighting. Additionally, the distant range sidelobes are associated with the ripples of the spectrum amplitude and, thus, cannot be removed by weighting. We show that by using a predistorted chirp with amplitude or phase shaping for amplitude ripple reduction and a correlation filter that accounts for the transducer's natural frequency weighting, output sidelobe levels of -35 to -40 dB are directly obtained. When an optimized filter is applied on receive, contrast or resolution can be traded in for range sidelobe levels down to -86 dB. The digital filter is designed to efficiently use the available bandwidth and at the same time to be insensitive to the transducer's impulse response. For evaluation of the method, simulations were performed with the program Field II.

A commercial scanner (B-K Medical 3535) was modified and interfaced to an arbitrary function generator along with an RF power amplifier (Ritec). Hydrophone measurements in water were done to establish excitation voltage and corresponding intensity levels (I_{sptp} and I_{spta}) well below the safety limits. In vivo images of the liver from two healthy volunteers show apparent increased depth of penetration of about 4 cm at 4 MHz, for codes of length 25 μs. Images taken from a non-attenuating wire phantom show that the -20 dB axial resolution for the coded scheme is as good as with pulse excitation (about 1.5λ), depending on the filter design criteria. The axial sidelobes are below -40 dB, which is the noise level of the measuring imaging system. The proposed excitation/compression scheme shows good overall performance and stability to the frequency shift due to attenuation. It increases the penetration depth, and still yields a high resolution and low adjustable sidelobes.

1 Introduction

Coded excitation in medical ultrasound can be used for improving the signal-to-noise ratio (SNR) and/or the penetration depth, as long as both sidelobe level and energy are kept acceptably low for ultrasound imaging. An improvement in SNR of 10 dB or more is expected [1], [2], depending on the code parameters, and the available bandwidth. As we reported previously [2], clinical images of the liver from two healthy volunteers demonstrated an improvement in penetration depth of at least 4 to 5 cm at 4 MHz with temporal sidelobe levels lower than -35 dB. In a high contrast imaging system this is not sufficient and this paper primarily focuses on an optimized design of pulse compression schemes applicable to ultrasound.

Linear FM waveforms have been used as coded excitation because of their important symmetry properties and their linear group delay. Non-linear chirps do not possess this linearity and therefore tend to be more sensitive to frequency shifts from the medium to be imaged. As it will be shown, linear chirps with amplitude tapering can give excellent performance, if they are designed properly. M-sequences and polyphase codes that have been used in radars are based on abrupt changes in phase that contain high harmonics, that the ultrasonic transducer smoothes out. Therefore, pulse compression with binary codes applied in ultrasound systems has worse performance than in radars due to the transducer, and have not been considered further.

Mismatched filtering

Range sidelobes represent an inherent part of the pulse compression mechanism, primarily due to the abrupt rise in the

signal spectrum. In radar systems, the most used approach for sidelobe reduction is weighting either in the time or in the frequency domain. These approaches are not equivalent for low TB products. Time weighting is preferable and has been used exclusively in this paper, because it yields lower sidelobe compression outputs [3]. Mismatched filtering has been treated extensively in the radar literature [4], [5]. Briefly, the transfer function of time weighting-based mismatched filters is the complex conjugate of an amplitude-weighted version of the excitation signal. Among a plethora of window functions, the Dolph-Chebyshev window exhibits the minimum main-lobe width for a specified sidelobe level and has been used extensively in radars [5]. We have used other windows such as the Blackman-Harris [6] windows, which yield narrow mainlobe width with adequately low sidelobe levels. The drawbacks for the low sidelobe level are a) a broadening of the mainlobe and associated loss in axial resolution and b) a loss in SNR.

The effect of the transducer

In ultrasound, bandpass filtering of the excitation signal occurs inherently from the transducer. The transducer effect on pulse compression is illustrated in Fig. 1, based on measurements of the pulse-echo impulse response. The transducer used is a 4 MHz single-element transducer (B-K Medical) and the calculated -6 dB fractional bandwidth was 65%. When the bandwidth of the chirp matches the transducer's bandwidth, the presence of the transducer reduces the near range sidelobes from -13.2 to -32 dB below the mainlobe. The linear FM signal used for Fig. 1 had 1.2 higher bandwidth, and that reduces the sidelobes further down to -42 dB. When mismatched filtering is applied (bold lines of Fig. 1), near range sidelobes can be further reduced at the expense of an additional broadening in the main lobe. The effect of the transducer is equivalent to an additional weighting that, although suboptimal, is beneficial to sidelobe reduction. In radars, weighting of the transmitted pulse has been avoided in order to send out more energy and all processing is done on the receiver.

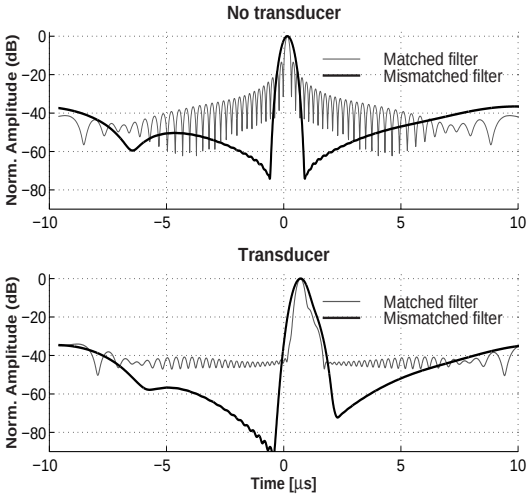

Figure 1: The effect of the transducer on pulse compression. The faint lines are the matched-filter outputs and the bold lines are the outputs when a Dolph-Chebyshev window has been applied to the compression filter. The specified sidelobe level for the window was -90 dB.

2 Reduction of the Fresnel ripples

As Fig. 1 shows, weighting severely fails to achieve the specified sidelobe levels in the ±T/2 region. This is an important consideration in the design of efficient coded excitation systems. The distant sidelobes increase the mainlobe-to-sidelobe energy ratio (MSR), an important image quality measure, which describes the leakage of energy from bright into dark areas of the image [7]. As it will be shown, the distant sidelobes are not related to the properties of the weighting function, but to the Fresnel ripples present in the spectrum of the FM signals. The efficiency of frequency windowing is based on the concept of shaping a rectangular spectrum to reduce the ripples of the sinc function in the time domain. For TB products larger than 50, the linear FM signal has approximately a rectangular envelope and a quadratic phase in both time and frequency domain, since the instantaneous frequency and group delay are linear functions of time and frequency respectively. This approximation is not valid for the low TB chirps that can be used in medical ultrasound. The exact solution for the spectrum is given by [8]:

$$S(f)=\frac{1}{\sqrt{2B}}e^{-j\pi f^2T/B}\{C(x_1)+C(x_2)\pm j\,[S(x_1)+S(x_2)]\} \quad (1)$$

where T is the duration of the FM signal, B its bandwidth, $D=TB$, $x_1=\sqrt{2D}\left(\frac{1}{2}-\frac{f}{B}\right)$ and $x_2=\sqrt{2D}\left(\frac{1}{2}+\frac{f}{B}\right)$. C and S are the quasi-oscillating Fresnel integrals that are responsible for the ripples in both spectrum amplitude and phase. The distant range sidelobes are paired-echo distortions due to deviation from the ideal rectangular spectrum amplitude and, thus, are not affected by any frequency weighting. This can be understood if the following Fourier pair is considered [5]:

$$\frac{a_n}{2}g\left(t+\frac{n}{B}\right)+g(t)+\frac{a_n}{2}g\left(t-\frac{n}{B}\right)\overset{\mathcal{F}}{\longleftrightarrow}G(f)\left[1+a_n\cos\left(2\pi n\frac{f}{B}\right)\right] \quad (2)$$

The presence of n spectrum amplitude ripples of amplitude a_n over the passband B of a signal spectrum $G(f)$ create sym-

metrical paired echoes in the time domain delayed and advanced from the main signal by n/B and scaled in amplitude by $a_n/2$. In a linear FM signal, the number of ripples and their amplitude depend on the TB product. In our case, the approximate paired-echo sidelobes are at about $\pm$T/2 relative to the center of the compressed pulse with an amplitude of about -30 to -40 dB [9]. The displayed dynamic range of a typical ultrasound image is 50 dB or greater, and therefore methods that compensate for the distant axial sidelobes have to be employed.

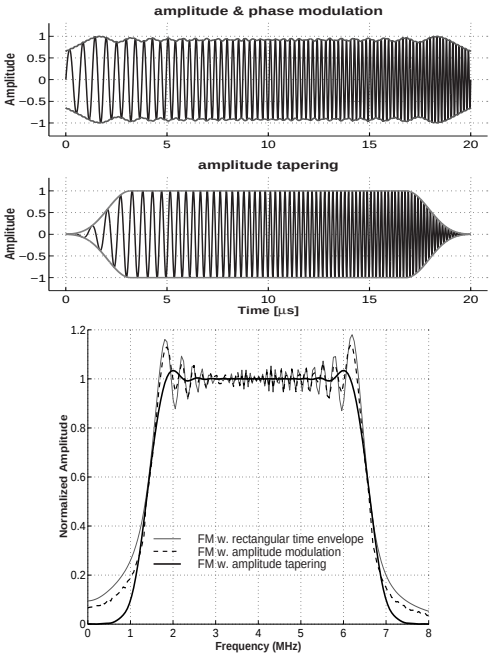

Figure 2: Reduction of spectrum ripples using envelope & phase modulation and amplitude tapering.

We have tested three different approaches for ripple reduction: envelope & phase modulation, phase predistortion, and time weighting of the edges.

Envelope modulation. This approach is based on the symmetry property of the linear FM signal in time and frequency. A linear FM signal with a constant amplitude envelope in the time domain yields a rectangular amplitude spectrum with ripple distortion. The reciprocal signal with a constant amplitude spectrum and strictly quadratic phase will have an amplitude and phase in the time domain given by Eq.(1), if T is replaced with B and t with f [8].

Phase predistortion. An alternative approach with similar performance is to use a phase predistortion of the transmitted signal, since amplitude and phase distortions have functional similarity and produce similar effects [9]. Both quadratic and cubic predistortion functions have been previously reported [9],[3]. Phase predistorted chirps have the advantage of sending out more energy, but are more vulnerable to phase distortions added from acoustic propagation.

Amplitude tapering. The amplitude ripples of the spectrum can be attributed to the sharp rise-time of the time envelope, since a linear FM signal with infinite duration has no ripples. Data correlating pulse rise-time and spectrum ripple have been previously reported [9]. Based on this analysis, a modified chirp with amplitude tapering of the transmitted signal has been generated. The attainable ripple reduction is a function of the signal's bandwidth, the choice of the tapering function and its duration. Fig. 2 shows the results of the different approaches on spectrum ripple reduction. The tapered function used in Fig. 2 is a Tukey window with a duration of 0.15T.

3 Waveform and filter design

As it was shown in the previous section, amplitude tapering is the most efficient way to reduce the Fresnel ripples of the spectrum, if the power amplifier allows control of the transmitted pulse rise time. For a given duration T, the design parameters are

i) the frequency band that is swept relative to the transducer's bandwidth

ii) the choice of the tapering function and its duration

iii) the choice of the weighting function.

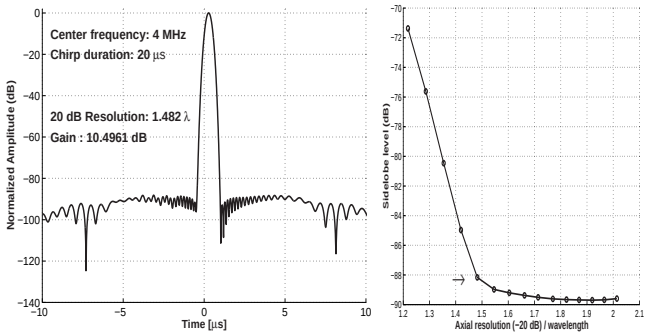

Figure 3: Optimized compression output. The right graph shows the tradeoffs between sidelobe level and axial resolution for a Dolph-Chebyshev mismatched compression filter. For most applications, the appropriate choice is at points in the knees of the curve as the one indicated by the arrow.

The transducer's bandwidth can be modeled, but the phase of its transfer function is unknown and, thus, can not be com-

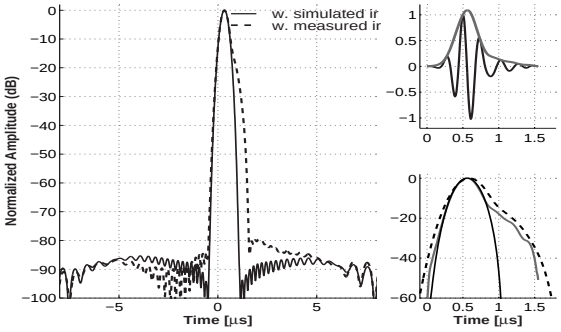

Figure 4: The effect of the transducer on the new scheme. The gray lines are the envelope of the measured impulse response and the dotted lines is the compressed output when the actual transducer is used. The black solid line in the bottom right figure is the logarithmic envelope of the simulated transducer.

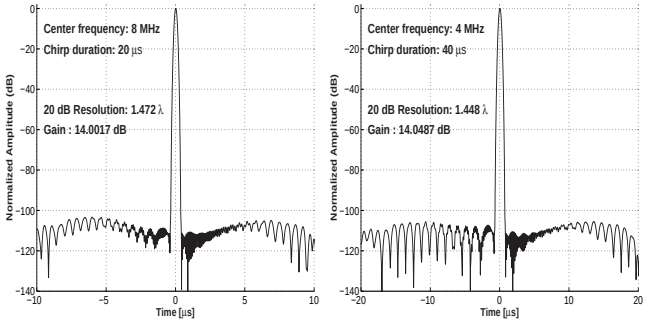

Figure 5: Optimized compression outputs for FM signals with amplitude tapering.

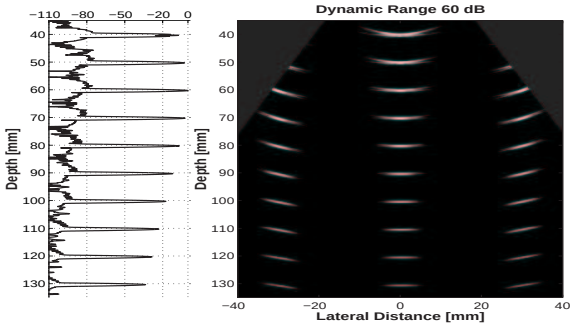

Figure 6: Simulated image of a phantom consisting of point scatterers in an attenuating medium with attenuation of 0.7 dB/[MHz×cm]. On the left of the graph, the central rf-line is plotted.

pensated for. The effort here is to minimize the effect of the convolution between the transducer impulse response and the excitation signal on the design of the waveforms and the compression filter. This is done in the excitation signal by sweeping a bandwidth slightly larger than the transducer's frequency passband. This has an additional advantage: the effect of the two main rise and fall overshoots of the spectrum (see Fig. 2) is minimized. With this design, the range sidelobes can be as low as -88.2 dB. Fig. 3 and Fig. 5 show optimized compression outputs for different frequencies and signal durations. The compression scheme is rather insensitive to the actual impulse response of the transducer (Fig. 4). The increased bandwidth of the excitation waveform also yields a better axial resolution and higher gain. The appropriate choice of the weighting function is a tradeoff between axial resolution and sidelobe levels (Fig. 3). If the length of the excitation signal is doubled, the sidelobes drop down to -105.5 dB with a superb resolution of 1.45λ (Fig. 5). For comparisons, the -20 dB axial resolution with a typical pulse excitation is 1.46λ. The design gives even better results for higher frequencies.

4 Simulated and experimental results

The new excitation/compression scheme has been tested with the simulation program Field II [10]. The simulated image of Fig. 6 shows good performance when attenuation is present.

An experimental coded excitation system has been used for the evaluation of the new method. A commercial scanner (B-K Medical 3535) was used along with a single-element mechanically rotating 4 MHz transducer. A detailed description and a block diagram of the system are given in [2]. The transmitter of the scanner was inactivated and the machine was modified to operate with external excitation. A LeCroy 9112 arbitrary function generator (AFG) was used for the generation of the coded waveforms on the computer. A power RF amplifier (Ritec 5000), specifically designed to drive ultrasound transducers, was used subsequently for amplification. The acquisition of the RF data was done using the sequential acquisition capabilities of an 8-bit digital oscilloscope (LeCroy LC334AM) at a minimum sampling sampling of 20 MHz depending on the requested depth of penetration and number of scan lines.

A linear FM signal 20 μs long was used, with a theoretical compression output as the one shown in Fig. 3. The power amplifier distorted the phase of the excitation signal, introducing spectral ripples. In order to evaluate the effect of this distortion, additional simulations were performed, in which

the actual output of the power amplifier has been used. To simulate the imaging system further, the measured impulse response of the transducer has been used. A phantom was scanned using the experimental coded system. The phantom consisted of wires placed every 1 cm in water. Additional wires were placed off-axis for evaluation of the system's resolution. The resulting image is shown in Fig. 7. For quantitative comparisons, the rf-lines are also shown. There is an apparent agreement between simulated and experimental results. Fig. 7 also shows how the SNR improves from pulse to coded excitation (top graphs).

Acquired clinical images are shown in Fig. 8.

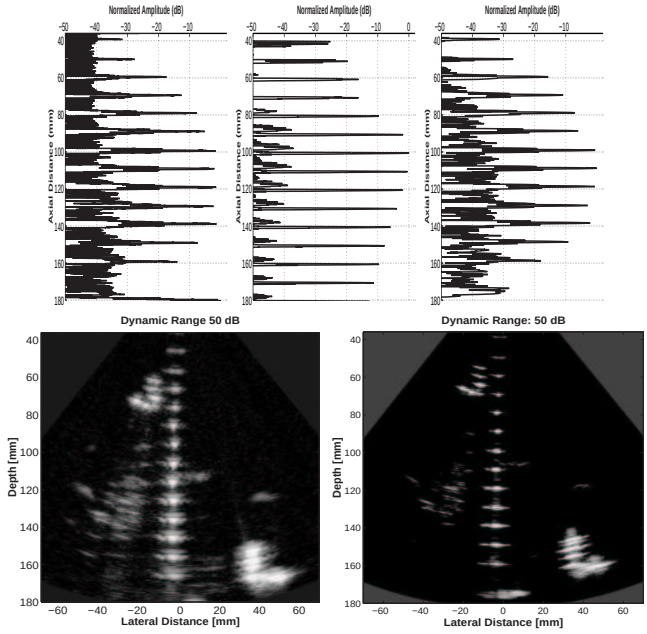

Figure 7: Acquired image of a wire phantom before and after compression. On the top of the graph, the central rf-lines are plotted from a) acquired image of the phantom with pulse excitation (not shown here), b) simulated image with coded excitation when the signal from the output of the power amplifier and the measured impulse response of the transducer have been used, and c) acquired image with coded excitation shown here.

5 Conclusions

We have presented a new coded excitation scheme with distinct features, that make it attractive to the implementation of high-performance coded ultrasound systems. The range resolution that can be achieved is comparable to that of a

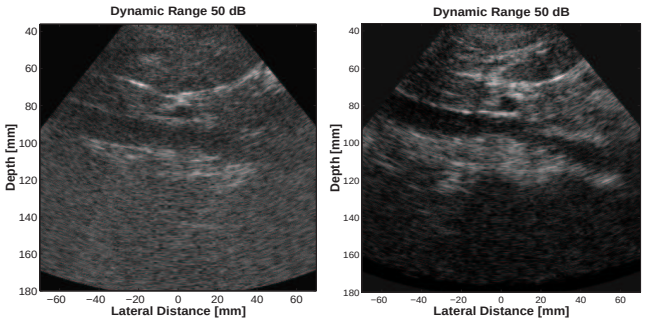

Figure 8: Acquired clinical images with pulse and coded excitation.

conventional system, depending on the available bandwidth. The range sidelobes, at the same time, are well-below the limits of the typical dynamic range of an ultrasound image. The energy of the sidelobe region is also reduced by lowering the distant sidelobes caused by the ripples of the spectrum's amplitude. Testing of the new coding excitation/compression scheme with Field II has shown good stability to the effect of frequency shift due to attenuation. Phantom images verified the good performance of the system. Compared to previously reported excitation systems, the proposed approach exhibits lower sidelobe energy (and therefore higher image contrast capabilities) and high compression gain, while maintaining axial resolution.

Acknowledgments

The authors would like to thank Kim Gammelmark, Christian H. Jørgensen, Niklas Lindberg, and Anders H. Thomsen for their work on the experimental setup, Dr. Morten H. Pedersen for some clinical scans and Ph.D. student Peter Munk for his valuable advices.

This work was supported by grant 9700883 and 9700563 from the Danish Science Foundation and by B-K Medical A/S.

References

[1] M. O'Donnell. Coded excitation system for improving the penetration of real-time phased-array imaging systems. *IEEE Trans. Ultrason., Ferroelec., Freq. Contr.*, 39:341–351, 1992.

[2] T. X. Misaridis, K. Gammelmark, C. H. Jørgensen, N. Lindberg, A. H. Thomsen, M. H. Pedersen, and J. A.

Jensen. Potential of coded excitation in medical ultrasound imaging. *Ultrasonics*, 38:183–189, 2000.

[3] M. Kowatsch and H. R. Stocker. Effect of Fresnel ripples on sidelobe suppression in low time-bandwidth product linear FM pulse compression. *IEE Proc.*, 129:41–44, 1982.

[4] C. E. Cook and M. Bernfeld. *Radar signals*. Artech House, Inc., 1993.

[5] M. I. Skolnik. *Radar Handbook*. McGraw-Hill, New York, second edition, 1990.

[6] F. J. Harris. On the use of windows for harmonic analysis with the discrete Fourier transform. *Proc. IEEE*, 66:51–83, 1978.

[7] J. W. Adams. A new optimal window. *IEEE Trans. Sig. Pro.*, 39:1753–1769, 1991.

[8] A. W. Rihaczek. *Principles of high-resolution radar*. McGraw-Hill, New York, 1969.

[9] C. E. Cook and J. Paolillo. A pulse compression predistortion function for efficient sidelobe reduction in a high-power radar. *Proc. IEEE*, 52:377–389, 1964.

[10] J. A. Jensen. Field: A program for simulating ultrasound systems. *Med. Biol. Eng. Comp.*, 10th Nordic-Baltic Conference on Biomedical Imaging, Vol. 4, Supplement 1, Part 1:351–353, 1996b.

A.3 Clinical use and evaluation of coded excitation in B-mode images

Paper presented at the IEEE International Ultrasonics Symposium, Puerto Rico, USA, 2000.

Clinical use and evaluation of coded excitation in B-mode images

Thanassis Misaridis[1], Morten H. Pedersen[2] and Jørgen Arendt Jensen[1]

[1]Center for Fast Ultrasound Imaging,
Department of Information Technology, Bldg. 344,
Technical University of Denmark,
DK-2800 Lyngby, Denmark

[2]Herlev University Hospital, Department of Ultrasound, DK-2730 Herlev, Denmark

Published in Proceedings of IEEE International Ultrasonics Symposium, Puerto Rico, USA, 2000.

Clinical use and evaluation of coded excitation in B-mode images

Thanassis X. Misaridis[1], Morten H. Pedersen[2] and Jørgen A. Jensen[1]

[1]Center for Fast Ultrasound Imaging,
Department of Information Technology, Bldg. 344,
Technical University of Denmark, DK-2800 Lyngby, Denmark
[2]Herlev University Hospital, Department of Ultrasound, DK-2730 Herlev, Denmark

Abstract

Use of long encoded waveforms can be advantageous in ultrasound imaging, as long as the pulse compression mechanism ensures low range sidelobes and preserves both axial resolution and contrast. A coded excitation/compression scheme was previously presented by our group, which is based on a predistorted FM excitation and a mismatched compression filter designed for medical ultrasonic applications. The attenuation effect, analyzed in this paper using the ambiguity function and simulations, dictated the choice of the coded waveform. In this study clinical images, images of wire phantoms, and digital video demonstrate the applicability, clinical relevance, and improvement attained with the proposed scheme.

A commercial scanner (B-K Medical 3535) was modified and interfaced to a software configurable transmitter board and to a sampling system with a 2 GB memory storage. The experimental system was programmed to allow alternating excitation on every second frame. That offers the possibility of direct comparison of the same set of image pairs; one with pulsed and one with encoded excitation. Abdominal clinical images from healthy volunteers were acquired and statistically analyzed by means of the auto-covariance matrix of the image data. The resolution laterally is retained, axially is improved, while there is a clear increase in penetration.

Phantom images using the proposed FM-based scheme as well as complementary Golay codes were also acquired, in order to quantitatively evaluate the characteristics of the compressed output and its stability to attenuation. Images of a wire phantom in water show that the range sidelobes resulting from pulse compression are below the acoustic noise, which was at 50 dB for the pulsed images. For images acquired from an attenuation phantom, the proposed compression scheme was robust to frequency shifts resulting from attenuation. The range resolution is improved 12% by the coded scheme compared to a 2-cycle pulse excitation. For the maximum acquisition depth of 15 cm, where the coded excitations are primarily intended, the improvement in SNR was more than 10 dB, while the resolution was retained.

1 Introduction

Application of coded excitation in medical ultrasound has been scarce in the literature [1],[2]. Most of these attempts have not paid much attention to the choice of coded waveform apart from its autocorrelation properties. However, in the highly attenuating scattering medium of the human body, it will be shown that the matched filter output is not the autocorrelation function, and attenuation should be taken into account in the coded signal design.

Modulated signals in attenuating media

Let $s(t) = \alpha(t) \cdot \cos[2\pi f_0 t + \phi(t)]$ be a real modulated signal transmitted by the ultrasound transducer, where f_0 is the center frequency of the transducer and $\alpha(t)$, $\phi(t)$ are the amplitude and phase modulation functions, respectively. If $s(t)$ is a narrowband signal, the complex representation $\psi(t)$ of $s(t)$ is analytical and can be written as

$$\psi(t) = \mu(t) \cdot e^{j2\pi f_0 t}, \tag{1}$$

where $\mu(t)$, often called the complex envelope of the signal, combines amplitude and phase modulation:

$$\mu(t) = |\mu(t)| \cdot e^{j\phi(t)}. \tag{2}$$

The real signal becomes:

$$s(t) = Re\{\psi(t)\} = |\mu(t)| \cos[2\pi f_0 t + \phi(t)], \tag{3}$$

When an ultrasound transducer is excited with such a modulated signal, the spectrum is weighted by the transducer bell-shaped spectrum. For a linear FM signal with a rectangular

spectrum amplitude, the shape of the transmitted pulse spectrum amplitude is effectively that of the transducer. When the medium has no attenuation, the returned signal from a single scatterer is simply a time-shifted version of the transmitted signal:

$$r(t) = \psi(t-\tau_0) = \mu(t-\tau_0)\cdot e^{j2\pi f_0(t-\tau_0)} \tag{4}$$

where τ_0 is the time instant after the start of transmission ($t=0$), at which the signal is being received.

The frequency dependence of the attenuation results in a larger attenuation of high frequencies than lower, when this signal propagates through the tissue. That will decrease the upper part of the bandwidth, causing effectively a decrease in the mean frequency that is a linear function of depth for a Gaussian-shaped pulse [3]. A significant downshift can be seen even for moderate tissue depths. For a transducer with a center frequency of 4 MHz and a 65% fractional bandwidth, there is a downshift of 800 KHz at a depth of 16 cm. For moving targets, there is an additional Doppler shift, which is, however, two orders of magnitude less than the frequency downshift due to attenuation.

Assuming that attenuation does not distort the complex envelope of the modulated signal, and it only causes a frequency downshift f_d, the returned signal in an attenuating medium will be

$$\begin{aligned} r(t) &\simeq \mu(t-\tau_0)\cdot e^{j2\pi[(f_0-f_d)(t-\tau_0)]} \\ &\simeq \mu(t-\tau_0)\cdot e^{j2\pi[f_0(t-\tau_0)-f_d(t-\tau_0)]}. \end{aligned} \tag{5}$$

Matched Filter Processing and the Ambiguity Function

The filter that maximizes the peak signal-to-noise ratio in the presence of white noise is the matched filter. The transfer function of the matched filter is the complex conjugate of the signal spectrum. Matched filtering of the received signal given in (5) requires *a priori* knowledge of the unknown parameters τ_0 and f_d. Practically there is no need to match to the target delay τ_0, since a change in delay merely changes the time at which the output occurs.

In the general case, the returned signal is matched to the signal $\psi(t-\tau)\cdot e^{j2\pi f_m(t-\tau)}$ with parameters f_m and τ. This corresponds to a filter with an impulse response of

$$h_{\tau,f_m}(t) = \psi^*(T-t+\tau)\cdot e^{-j2\pi f_m(T-t+\tau)}. \tag{6}$$

The receiver's output at time T (translated so that the maximum occurs at $\tau=0$) will then be

$$\psi_{out}(\tau) = e^{j2\pi f_0\tau_0}\cdot\chi(\tau-\tau_0, f_m-f_d), \tag{7}$$

where $\chi(\tau, f_d)$ is defined as the *ambiguity function*, and is given by

$$\begin{aligned} \chi(\tau, f_d) &= \int_{-\infty}^{\infty} \psi(t)\cdot\psi^*(t-\tau)\cdot e^{j2\pi f_d t}dt \\ &= \int_{-\infty}^{\infty} \mu(t)\cdot\mu^*(t-\tau)\cdot e^{j2\pi f_d t}dt. \end{aligned} \tag{8}$$

The ambiguity function (AF) is the complete 2-D matched-filter response in scatterer range and frequency downshift. In the absence of attenuation (or generally when $f_m - f_d = 0$), the matched filter output is the autocorrelation function of the waveform.

The squared modulus of the ambiguity function is referred to as the *ambiguity surface* (AS) [4]. The volume under the AS indicates the total interference contributed by each target, and the distribution of the volume is the distribution of this interference in range and frequency shift [4]. The volume under the AS is only a function of the signal's energy and cannot be reduced by any modulation. Coding simply redistributes this volume on the $\tau - f_d$ plane.

Implications of attenuation on signal design

Tissue attenuation in ultrasound causes a significant frequency shift of the received signal. For coded excitation and some kind of matched filtering on the receiver, the compressed output is not the autocorrelation function of the code, but a cut of the ambiguity function for a certain frequency downshift. Therefore the AF of the transmitted waveform ought to have desired properties in the entire delay-frequency shift plane. One solution is to use a filter bank

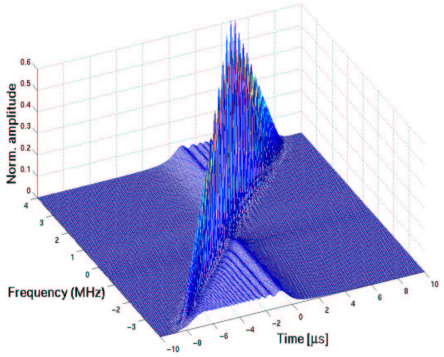

Figure 1: Ambiguity function of the linear FM waveform.

of i depth-dependent filters with impulse responses $h_{\tau,f_{m_i}}$, so that $f_{m_i} = f_{d_i}$, where f_{d_i} is the estimated frequency shift at a certain depth. However, depending on the interrogating

medium, there will always be deviations from the estimated values, and the requirement for robustness of compression in frequency shifts remains.

If attenuation is not compensated for, the ambiguity function should not have increased sidelobes off the delay axis. Figure 1 shows the ambiguity function of the proposed linear FM waveform. The advantage of this signal is that compression still occurs for a frequency mismatch. This occurs because its group delay is a linear function of frequency, and thus, the effect of frequency shift is a simple time translation of the compressed pulse. Note that a coded waveform with a nonlinear phase function (for example a non-linear FM waveform) will not be compressed, when there is a frequency shift.

Most binary phase codes share the undesirable property of being quite sensitive to frequency shifts. The AF of such signals will generally have a central spike on a wide pedestal of range-frequency sidelobes. A search for binary codes with good autocorrelation functions is quite useless in ultrasound, if the ambiguity function away from the delay axis is unsuitable [4]. The Barker codes, for instance, do not retain their ”perfect” compression output even with moderate frequency shifts. Binary phase codes insensitive to frequency shifts do exist, as for instance the Frank polyphase code, which is a discrete approximation of contiguous linear FM segments. Thus, the shape of the ambiguity function should be an additional design tool for search of such codes.

Complementary codes consist of two or more sequences of equal length, whose autocorrelation functions have sidelobes

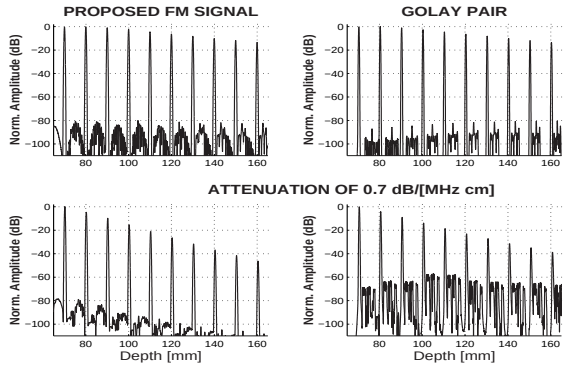

Figure 2: The effect of attenuation in the complementarity properties of Golay codes. Simulation results show a comparison of the echoes from 10 point scatterers between the proposed FM-based compression scheme and a Golay pair. The upper part is when there is no attenuation in the medium.

equal in magnitude but with inverse sign. Adding the autocorrelations of such pairs cancels the sidelobes entirely and doubles the mainlobe. The most well know complementary codes are the Golay codes. Simulation results in Fig. 2 show a comparison of the echoes from 10 point scatterers for the proposed FM-based compression scheme and a Golay pair. These plots show how these coded waveforms are affected by attenuation. While perfect sidelobe cancellation occurs for the complementary Golay codes, when the medium has no attenuation (upper plots), the complementarity property degrades with attenuation. This can be assessed by inspecting the ambiguity function of Golay codes. With an attenuation of 0.7 dB/[MHz$\times$cm], the range sidelobes increase up to -25 dB at a depth of 16 cm, in contrast to the proposed tapered FM signal, whose compression is very robust to attenuation.

Based on these observations, the linear FM waveform seems the appropriate coded waveform for the highly attenuating medium of medical ultrasound, as long as the pulse compression mechanism ensures low range sidelobes and good axial resolution. Techniques for linear FM signal predistortion and mismatched filtering have been presented previously by our group [5]. In this paper, we evaluate the performance of these techniques experimentally. Images with complementary phase codes are also evaluated in an attenuating phantom.

2 Materials and Methods

Measurement System

An experimental digital coded excitation system was constructed (Fig. 3). A commercial scanner (B-K Medical 3535) was interfaced to one transmitter and one receiver board from the Center's newly constructed RASMUS experimental system [6]. A single-element mechanically rotating 4 MHz

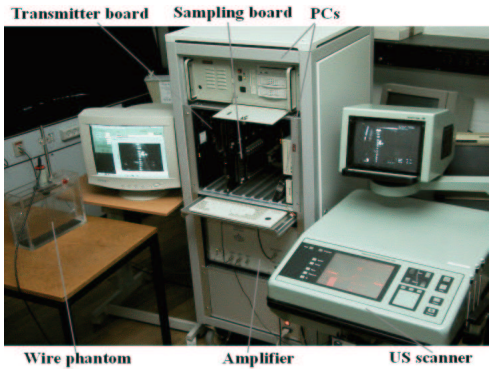

Figure 3: The experimental system.

probe (B-K Medical 8536) was used. The software control of both boards is implemented as a MATLAB^{TM} toolbox of

high-level commands. The transmitter board is capable of transmitting different complex arbitrary waveforms for each line of an image with a few lines of MATLAB^{TM} code. The board was programmed to allow alternating excitation on every second frame. That allowed direct comparison of the same set of image pairs; one with conventional and one with encoded excitation, or image pairs with complementary sequences. The analog RF data of the scanner were sampled at 40 MHz. The receiver sampling unit has 12-bit ADCs and 2 GBytes storage SDRAM. That allowed storage of 140 consecutive ultrasound images, corresponding to about 25 s of scan data. Subsequently the stored RF data were read from the memory, and all post processing (pulse compression, interpolation, scan conversion and display) was done off-line on the computer.

Performance evaluation

The performance of a pulse compression system can be measured using the four following indices:

i) The range mainlobe width, a measure of the range resolution.

ii) Range sidelobe measures, a measure of the induced "self-noise". Frequently used indices are the peak sidelobe level (PSL) and the integrated sidelobe level (ISL). PSL is simply the ratio between the highest sidelobe and the mainlobe. ISL is the the ratio of the total sidelobe energy to main peak energy expressed in decibels.

iii) The gain in signal-to-noise ratio (GSNR).

iv) Robustness to attenuation. There is no quantitative index reported. The shape of the ambiguity function can be used as a qualitative tool.

We have reported a compression scheme using a predistorted linear FM waveform and weighted mismatched filtering that has excellent measures of these indices [5]. The compressed output from this scheme has a theoretical -20 dB mainlobe width of 1.48 λ, very low axial sidelobes (PSL=88 dB) and GSNR=10.5 dB. Figure 1 shows the robustness of this waveform to ultrasound attenuation.

3 Experimental results

Initially, pulse-echo measurements from a wire phantom in water were performed. It has not been possible to evaluate the sidelobe level of the compressed output for the proposed code, but we can conclude that the sidelobes are below -50 dB, which was the acoustic noise for the pulsed excitation output. The acoustic noise was due to reverberation at the water surface and the bottom of the tank. The -6 dB axial

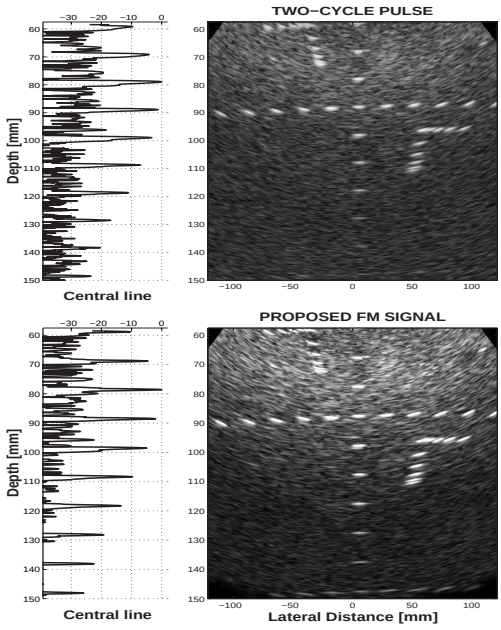

Figure 4: Detail of images of a wire phantom (right) and central RF lines (left) for coded and pulsed excitation. Matched filtering has been applied to both images. The dynamic range of the images is 45 dB. There is an improvement in SNR of more than 10 dB. Axial resolution is also higher for the coded image.

resolutions were measured and were found to be 0.94λ for the coded and 1.18λ for a two-cycle pulsed excitation.

Images of an attenuating wire phantom are shown in Fig. 4 and 5. The power levels of the excitation signals on the transducer were lower than the original excitation power of the scanner. The phantom has attenuation of 0.5 dB/[MHz$\times$cm] and consists of wires of 0.2 mm in in diameter, positioned every 1 cm axially. Additional wires are placed at a 15 degree angle with decreasing distance down to 1 mm. In the pulsed image of Fig. 4 a two-cycle pulse of the carrier frequency was used. In the coded image the excitation was a 20 μs tapered FM signal with the compression filter being the one from [5]. The -6 dB axial resolution of the coded image measured at the wires in depth 12 and 14 cm was 1.31λ. The resolution of the pulsed image was 1.49λ. In conventional imaging, short broadband pulses have to be used in order to achieve good axial resolution and to use all available system bandwidth, with the drawback of degrading the SNR. When FM signals are transmitted, the axial resolution depend on the bandwidth that is swept and not on the duration of the signal. Therefore, coded images using FM modulated signals result not only in

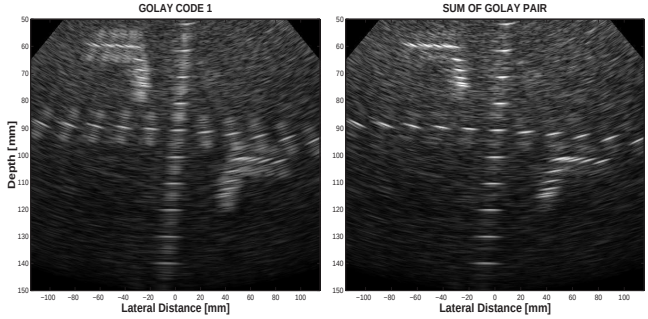

Figure 5: Images with Golay pair excitation of a wire phantom with attenuation of 0.5 dB/[MHz×cm]. On the left is the image with one of the Golay codes and on the right is the sum of the two complementary images. Dynamic range=45 dB.

a better SNR, but also in better axial resolution by utilizing the available bandwidth more efficiently. Fig. 4 shows that there is an effective gain in SNR of 10 dB or more, that corresponds to an additional penetration of 3-4 cm with the 4 MHz probe.

Images of the same phantom using a complementary pair of Golay codes with length 40 are shown in Fig. 5. A single Golay code has high axial sidelobes that are visible on the left image. When the echoes from 2 Golay codes are added coherently, there is a degree of cancellation on the axial sidelobes. However, confirming our hypothesis, the cancellation of the sidelobes is not perfect due to attenuation, and shadows are still visible along the wires.

Clinical images of the abdomen using the proposed scheme are shown in Fig. 6. The images show an excellent performance of the encoded excitation in terms of noise reduction at large depths and resolution. The auto-covariance matrix on the image gives an indication of the speckle size. The lateral resolution of speckle data from the images of Fig. 6 is very similar for both images, while axially the coded image has slightly better performance (Fig. 7).

These results clearly demonstrate that abdominal ultrasound imaging can benefit from coded excitation yielding a higher SNR and therefore penetration, while maintaining both axial and lateral resolution. The higher SNR can be exchanged with resolution by increasing the center frequency, i.e. for GSNR=10 dB, going from 4 to 5 MHz without compromising SNR. Longer codes can make this frequency step even bigger.

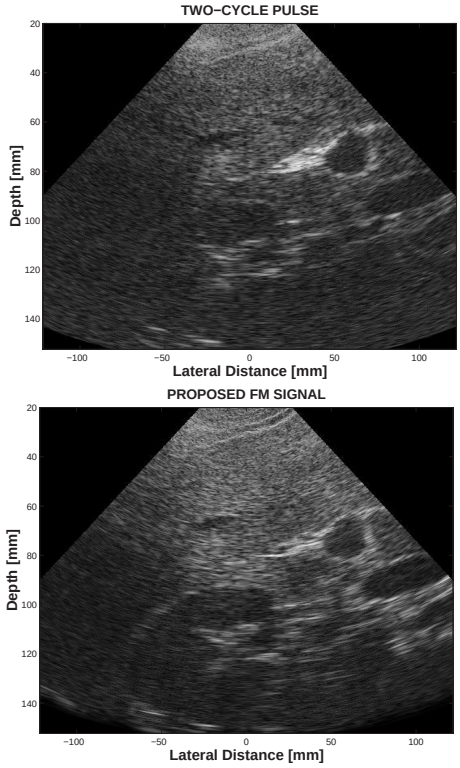

Figure 6: Clinical images of the right kidney for coded and pulsed excitation. The portal vein and the inferior vena cava are at the right side of the images and liver tissue is left from the kidney. The dynamic range of the images is 45 dB. Improvement in resolution and noise reduction at large depths are visible.

Acknowledgments

The authors would like to thank Ph.D. student B. Tomov for his extensive work on the experimental system and Dr. P. Munk for valuable discussions. This work was supported by grant 9700883 and 9700563 from the Danish Science Foundation and by B-K Medical A/S.

References

[1] M. O'Donnell. Coded excitation system for improving the penetration of real-time phased-array imaging systems. *IEEE Trans. Ultrason., Ferroelec., Freq. Contr.*, 39:341–351, 1992.

[2] J. Shen and E. S. Ebbini. A new coded-excitation Ultrasound imaging system - Part I: Basic principles. *IEEE*

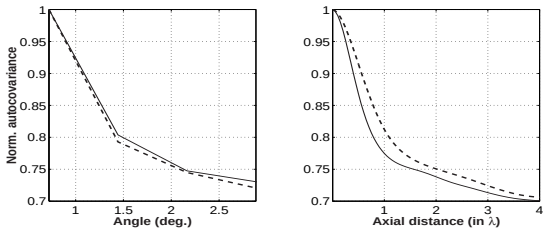

Figure 7: Lateral and axial resolution in speckle using autocovariance matrix analysis. The dashed lines are for the pulsed image.

Trans. Ultrason., Ferroelec., Freq. Contr., 43:131–140, 1996.

[3] J. A. Jensen. *Estimation of Blood Velocities Using Ultrasound: A Signal Processing Approach*. Cambridge University Press, New York, 1996.

[4] A. W. Rihaczek. *Principles of high-resolution radar*. McGraw-Hill, New York, 1969.

[5] T. X. Misaridis and J. A. Jensen. An effective coded excitation scheme based on a predistorted FM signal and an optimized digital filter. In *Proc. IEEE Ultrason. Symp.*, volume 2, pages 1589–1593, 1999.

[6] J. A. Jensen, O. Holm, L. J. Jensen, H. Bendsen, H. M. Pedersen, K. Salomonsen, J. Hansen, and S. Nikolov. Experimental ultrasound system for real-time synthetic imaging. In *Proc. IEEE Ultrason. Symp.*, volume 2, pages 1595–1599, 1999.

A.4 Space-Time Encoding for High Frame Rate Ultrasound Imaging

Space-Time Encoding for High Frame Rate Ultrasound Imaging

Thanassis Misaridis and Jørgen Arendt Jensen

Center for Fast Ultrasound Imaging,
Ørsted-DTU, Bldg. 348,
Technical University of Denmark,
DK-2800 Lyngby, Denmark

To be published in Ultrasonics, 2002.

Space-Time Encoding for High Frame Rate Ultrasound Imaging

Thanassis X. Misaridis and Jørgen A. Jensen

Center for Fast Ultrasound Imaging,
Ørsted-DTU, Bldg. 348,
Technical University of Denmark, DK-2800 Kgs.Lyngby, Denmark

Abstract

Frame rate in ultrasound imaging can be dramatically increased by using sparse synthetic transmit aperture (STA) beamforming techniques. The two main drawbacks of the method are the low signal-to-noise ratio (SNR) and the motion artifacts, that degrade the image quality.

In this paper we propose a spatio-temporal encoding for STA imaging based on simultaneous transmission of two quasi-orthogonal tapered linear FM signals. The excitation signals are an up- and a down-chirp with frequency division and a cross-talk of -55 dB. The received signals are first cross-correlated with the appropriate code, then spatially decoded and finally beamformed for each code, yielding two images per emission. The spatial encoding is a Hadamard encoding previously suggested by Chiao et al. [1]. The Hadamard matrix has half the size of the transmit element groups, due to the orthogonality of the temporal encoded wavefronts. Thus, with this method, the frame rate is doubled compared to previous systems. Another advantage is the utilization of temporal codes more robust to attenuation.

With the proposed technique it is possible to obtain images dynamically focused in both transmit and receive with only 2 firings. This reduces the problem of motion artifacts. The method has been tested with extensive simulations using Field II. Resolution and SNR are compared with uncoded STA imaging and conventional phased array imaging. The range resolution remains the same for coded STA imaging with 4 emissions and is slightly degraded for STA imaging with 2 emissions due to the -55 dB cross-talk between the signals. The additional proposed temporal encoding adds more than 15 dB on the SNR gain, yielding a SNR at the same order as in phased array imaging.

1 Introduction

Phased array imaging, currently used in ultrasound scanners, involves transmission of pulses from the entire transmit aperture, which are relatively delayed to form a focused beam along a given direction. The echoes received by all elements are beamformed to yield the image points along this direction, and the procedure is repeated for all image directions. There are two main limitations of the method: i) the acquisition time is proportional to the number of lines in the image, ii) the image has a fixed transmit focus.

Synthetic transmit aperture (STA) imaging can overcome both these problems. In STA imaging, one element transmits a pulse and all elements receive the echoes. Since each transmission is a spherical wave insonifying the entire imaging region, receive beamforming for all lines can yield a whole low-energy image for every transmission. Then the next element is excited and this is repeated until the whole transmit aperture is synthesized. The final synthetic image is the coherent sum of all beamformed images [2]. In this way, the final image is fully focused both in transmit and receive.

The frame rate in STA imaging does not depend on the number of scan lines, but on the number of transmit elements. If only two to five elements are used for transmitting, a frame rate on the order of 1,000 frames/s is possible, which will pave the way to 3-D imaging. The use of such small number of transmit elements is dictated by the problem of artifacts caused by tissue or transducer motion. In contrast to phased array imaging, where each line is formed from a single transmission, in STA imaging each line of the final image uses data from all transmit events, which sets phase coherence requirements. Using 5 transmissions with an 8 MHz transducer and a corresponding wavelength of about 190 μm, tissue speed greater than 14 mm/s may cause motion phase artifacts.

Such sparse transmit aperture in STA imaging can yield a high frame rate, no motion artifacts, and potentially optimally-focused images, but rises two new issues. The first is the resulting grating lobes in the radiation pattern of the array, since the necessitated sparse transmit aperture violates the half-wavelength element spacing. This problem has been discussed in the literature [2, 3] and will not be addressed here. It can be solved using the effective aperture concept. Briefly, the convolution of the transmit with the receive aperture for each transmission is the effective sub-aperture. The two-way radiation pattern of the synthetic image is the Fourier transform (at least in the far field) of the sum of the effective sub-apertures of all transmit events. Therefore, by proper apodization of the receive aperture for each transmit event, a uniform $\lambda/2$ spacing effective aperture can be created, even in the case of only 2 transmit events [3].

The second problem of sparse STA imaging compared to phased-array imaging is the low signal-to-noise ratio (SNR). This paper suggests a combined spatio-temporal coding based on linear FM signals that yields images with comparable SNR to conventional imaging.

2 Methods for increasing the SNR in STA imaging

It has been reported [4] that the SNR for a non-sparse STA imaging system, which uses all transmit elements N, is $20\log(\sqrt{N})$ less than the SNR of an equivalent phased-array imaging system. This is because the power transmitted by all elements in phased-array imaging is at the focal point N times higher than the power transmitted by 1 element as in STA imaging. Then, for a 64-element system, there should be a degradation in SNR of $10\log 64 = 18$ dB. For a STA system with 4 transmissions, the expected SNR degradation will then be at the order of 30 dB.

There have been three methods reported in the literature for increasing the SNR in STA imaging. In the first [4], transmit defocusing delays are applied to a group of elements, in order to obtain a similar one-way radiation pattern as a single transmitting element at the center of the group. In this case, the mean intensity of the emulated spherical waves increases by n (and not n^2 due to defocusing), as demonstrated experimentally by Karaman et. al. [4], where n is the number of simultaneously transmitted elements. For 11 elements an improvement in SNR of about $10\log 11 = 10.5$ dB should be expected.

In the second method [1], all N elements simultaneously transmit spherical waves with excitation pulses spatially encoded by the rows of a $N \times N$ matrix, such as the Hadamard matrix. After N spatially-encoded transmit events, the received data sets can be decoded, by the inverse matrix, yielding the equivalent N data sets as if each of the N elements was transmitting alone in every transmit event. The advantages of the Hadamard matrix over other orthogonal matrices are: i) it only consists of 1 and -1 (making coding implementation easy as a phase reversal of the transmitting pulse), and ii) the inverse Hadamard matrix is a scaled version of itself. Hadamard spatial encoding increases SNR by $10\log M$ [1], where M is the number of simultaneous transmitting elements (or group of elements). For a 4×4 Hadamard matrix considered here, there is an improvement of 6 dB.

The third reported method combines coded excitation with spatial Hadamard coding [5]. For four transmit events, it uses two pairs of complementary codes, while the four codes are also mutually orthogonal. The decoding process is based on the Hadamard matrix of half length, assisted by the cancellations of the sidelobes in the sum of complementary auto-correlations and cancellations of the sum of cross-correlations among mutual orthogonal codes. The improvement in the SNR is $10\log(M \cdot TB)$, where TB is the time-bandwidth product of the codes, which is equal to the length of the code L for binary codes. Using four transmit events and codes of length 32, as reported in [5], there is a $10\log 4 \cdot 32 = 21$ dB improvement.

3 Spatio-temporal encoding using linear FM excitation

The encoding scheme proposed here is a modification of the aforementioned method. It is apparent that the SNR can be significantly increased if long coded sequences are used instead of short pulses. It has been previously shown [6], that for the highly attenuating medium of medical ultrasound the linear FM waveform is the most appropriate coded waveform, as long as the pulse compression mechanism ensures low range sidelobes and good axial resolution.

A 64-element linear array with $\lambda/2$ element spacing has been used in the simulations. The entire array was used in receive for all transmit events with the same Hanning window for apodization of the received aperture. Such apodization

does not give the optimal effective aperture, and therefore higher lateral sidelobes compared to phased-array imaging are expected. For each of the 4 transmit events, the elements 1, 22, 43, 64 transmit the same tapered linear FM waveform. The design of the tapered FM signal and the weighted compression filter, which are used, have been described in [7]. The received signal, which has been convolved with the two-way transducer impulse response has an effective time-bandwidth product of 130. The compressed output from this compression scheme has a theoretical -20 dB mainlobe width of 1.48 λ and very low axial sidelobes (mainlobe-to-peak sidelobe ratio of -88 dB).

Upon reception, each of the 4 rf-data sets is decoded by the Hadamard matrix. The decoded rf-data are subsequently compressed using the same compression filter in all channels. Finally, the decoded compressed rf-data are beamformed for all directions, yielding a low-energy image. Fig. 1 shows simulated encoded STA images before and after compression obtained with the simulation program Field II [8]. Fig. 2 shows that the resolution both laterally and axially remains

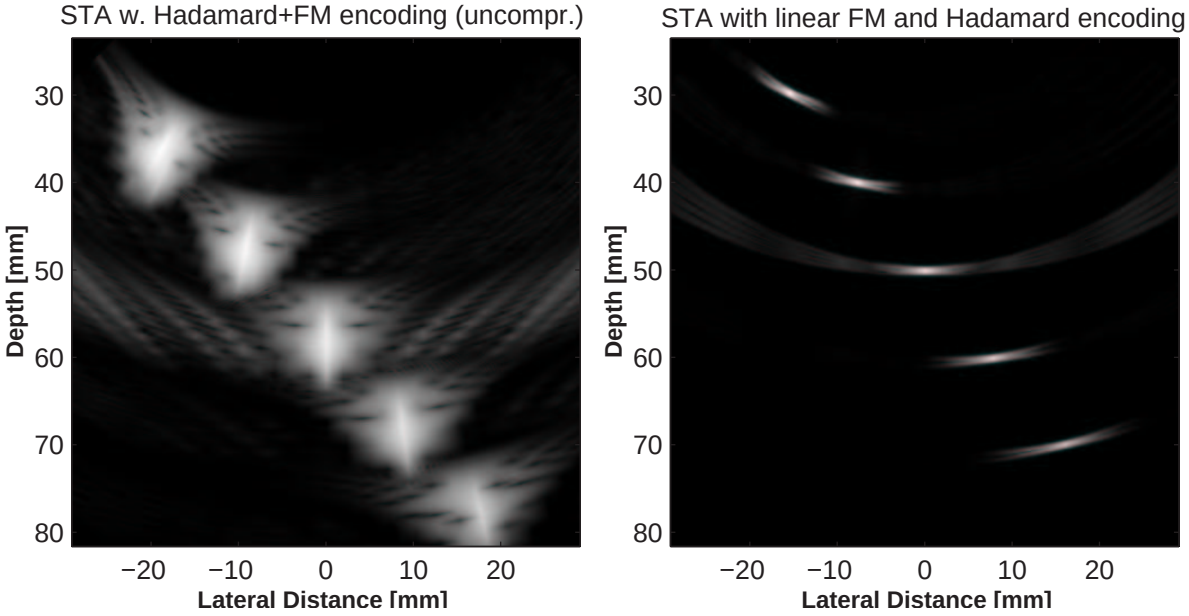

Figure 1: Simulated image of point targets for STA imaging using Hadamard encoding and tapered linear FM signals, before (left) and after compression (right). The dynamic range of both images is 60 dB.

practically the same, when Hadamard spatial encoding combined with FM time encoding is used. Also notice that the matched filter aligns the rf-data axially at the center of each scatterer, in contrast to pulsed excitation where each scatterer appears a little further than it actually is, due to convolution with the transducer impulse response.

This coding strategy yields the same frame rate and approximately the same SNR improvement as with the one reported in [5]. The advantage of the proposed method lies in the robustness of compression of linear FM signals, and the immunity this signal has in frequency shifts. On the other hand, the complementarity of Golay codes breaks down at large depths due to attenuation in tissues, resulting in image degradation. It has been shown [6] that with an attenuation of 0.7 dB/[MHz$\times$cm], the range sidelobes for a Golay pair increase up to -25 dB at a depth of 16 cm, in contrast to the proposed tapered FM signal, whose compression is very robust to attenuation. The temporal FM coding can also work independently in the absence of the spatial Hadamard coding, if the receiver has to be kept simpler. This is not the case for the complementary coding suggested in [5], where it is the Hadamard coding which cancels the sum of the cross-correlations between the mutual orthogonal codes. Finally, in the proposed method, image quality can be traded off with doubling of the frame rate.

This can be done by transmitting two linear FM signals from two pairs of transmitted elements, as shown in Table 1. If the two signals have low cross-correlation properties, a 2×2 Hadamard matrix can decode the four signals. In this way, two images can be beamformed for every transmit event and the frame rate is doubled at a small cost of resolution and

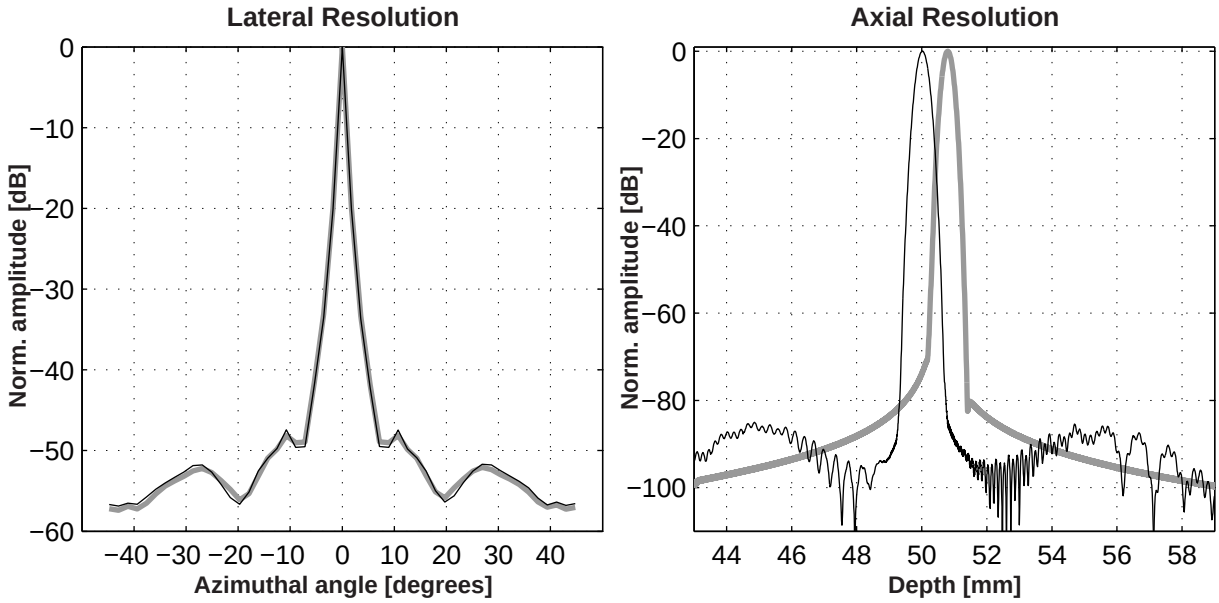

Figure 2: Lateral and axial resolution calculated from the simulated images at the point at depth 50 mm. The gray lines correspond to the typical STA image with 4 emissions, and the black lines to the STA image with the proposed Hadamard+FM encoding.

Table 1: Excitation scheme for fast coded STA imaging using two orthogonal FM signals C_1 and C_2.

TRANSMIT	ELEMENT			
EVENT	1	22	43	64
1	C_1	C_1	C_2	C_2
2	C_1	-C_1	C_2	-C_2

SNR. The two excitation signals are an up- and a down-chirp with frequency division. The auto- and cross-correlation of the two signals are shown in Fig. 3a. The results shown in Fig. 3 are based on simulation of a transducer at 8 MHz with a 65% fractional bandwidth and two broadband preweighted FM signals centered at 6 and 10 MHz. A small loss in SNR is expected since the transmitted energy is not centered at the transducer central frequency. The different FM slopes allow separation of about 12 dB between the two signals. Frequency division is responsible for further cross-talk reduction. The resulting STA image using only 2 emissions is shown in Fig. 3b.

4 SNR improvement

Fig. 4 shows simulated images of 5 point targets using the various coding schemes. White Gaussian noise was added to all the received channels before beamforming, to simulate electronic receiver noise. Equal driving voltages are used for all images. Note, however, that phased array imaging systems focus the pulses of all elements at the transmit focusing point. In STA imaging, only a single element transmits an unfocused wave. Therefore, for equal driving voltages, the transmitted power in STA imaging is much lower, and thus, the transmitted pulse amplitude can be increased more than an order of amplitude compared to conventional imaging without exceeding the intensity limitations [2]. It is, thus, possible to achieve up to 10 dB higher SNR than what is shown here, as long as the transducer elements can deliver such high power.

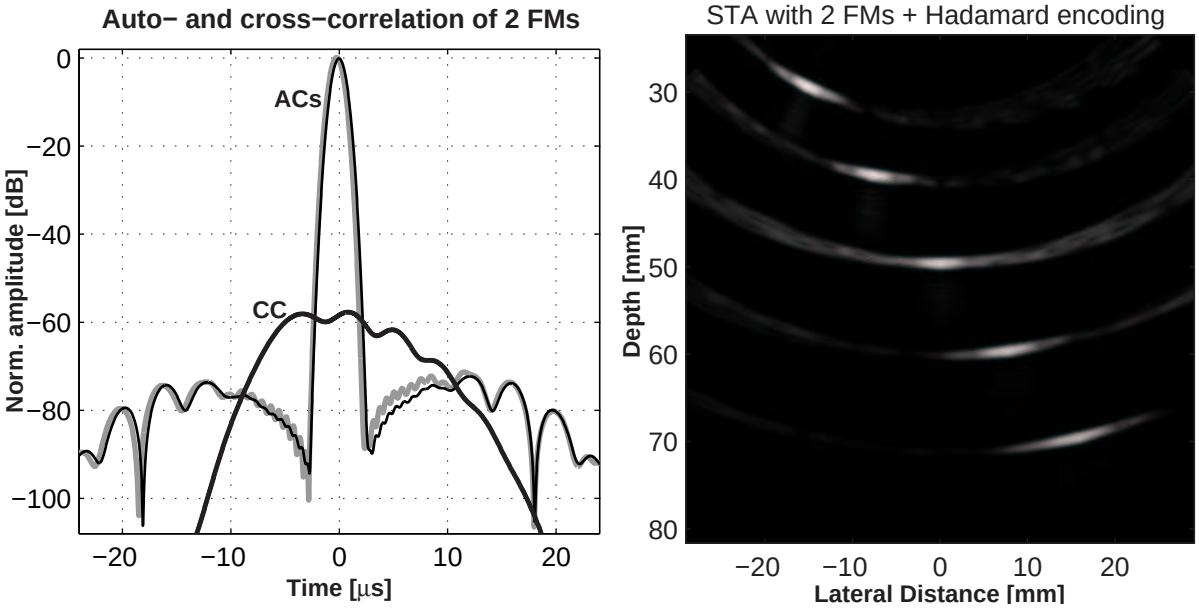

Figure 3: STA simulated image (right) with double frame-rate using Hadamard encoding and two orthogonal preweighted linear FM signals. The dynamic range of the image is 60 dB. On the left, the auto- and cross-correlation of the two signals are shown.

The gain in SNR that can be achieved by using spatially encoded FM excitation is apparent from the images of Fig. 4. For quantitative evaluation, the SNR has been estimated from the central line of the images. The results are listed in Table 2. The entry for the phased array image is the maximum SNR, calculated at the transmit focal point. In the case of STA

Table 2: Calculated SNR from the central line of the simulated images

IMAGING MODE	EMISSIONS	SNR (dB)	DEPENDENCE
Phased Array Imaging	51	53.09	$10\log(64\cdot 64)$
STA w. 1 xmit event	1	18.64	0
STA w. 64 xmit events	64	36.43	$10\log(64)$
STA w. 4 xmit events	4	24.34	$10\log(4)$
STA w. 4×4 Hadamard coding	4	30.33	$10\log(4\cdot 4)$
STA w. 4×4 Hadamard + linear FM	4	52.74	$10\log(4\cdot 4\cdot TB)$
STA w. 2×2 Hadamard + 2 linear FMs	2	47.56	$10\log(2\cdot 2\cdot(TB_1+TB_2))$

with 2 coded transmissions using orthogonal chirps, there is a trade off between low cross-correlation among the signals and loss in SNR. For a different design of the two codes than the one presented here, cross-correlation around -65 dB is possible with an additional loss of 5 dB in the SNR. In conclusion, the suggested coding scheme can yield synthetic aperture images with four transmissions, which can have the same SNR as that for phased-array imaging at the transmitted focal point, while retaining image resolution. In combination with the method of defocused transmit subapertures, this SNR can be achieved with only two transmissions of orthogonal preweighted linear FM signals.

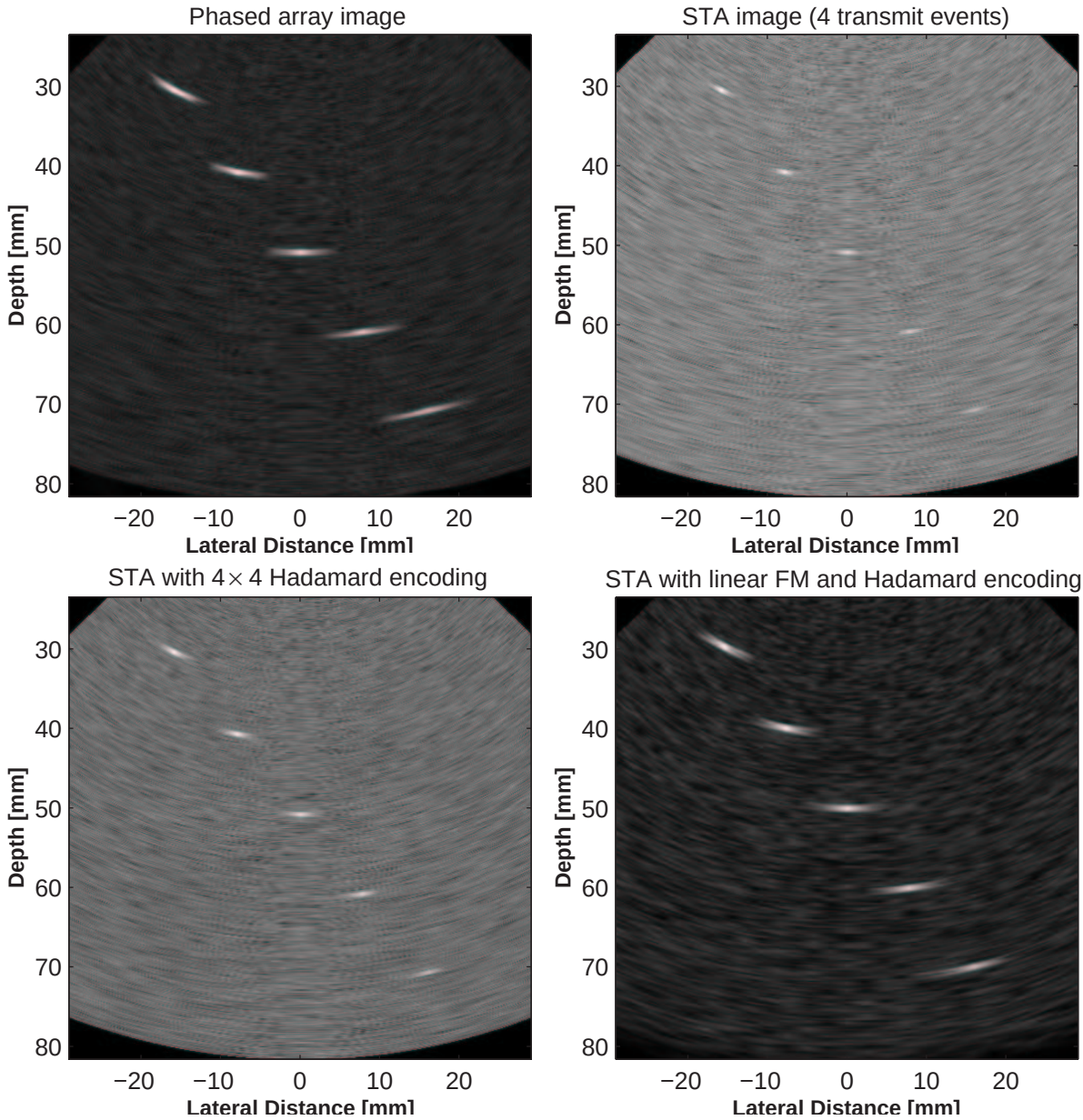

Figure 4: Simulated images of point targets with added noise, showing the improvement in SNR for the various coding schemes. The dynamic range of all images is 60 dB.

References

[1] R. Y. Chiao, L. J. Thomas, and S. D. Silverstein. Sparse array imaging with spatially-encoded transmits. In *Proc. IEEE Ultrason. Symp.*, pages 1679–1682, 1997.

[2] G. R. Lockwood, J. R. Talman, and S. S. Brunke. Real-time 3-D ultrasound imaging using sparse synthetic aperture beamforming. *IEEE Trans. Ultrason., Ferroelec., Freq. Contr.*, 45:980–988, 1998.

[3] C. R. Cooley and B. S. Robinson. Synthetic aperture imaging using partial datasets. In *Proc. IEEE Ultrason. Symp.*, pages 1539–1542, 1994.

[4] M. Karaman, P. C. Li, and M. O'Donnel. Synthetic aperture imaging for small scale systems. *IEEE Trans. Ultrason., Ferroelec., Freq. Contr.*, 42:429–442, 1995.

[5] R. Y. Chiao and L. J. Thomas. Method and apparatus for ultrasonic synthetic transmit aperture imaging using orthogonal complementary codes. *US Patent*, 6,048,315, 2000.

[6] T. X. Misaridis, M. H. Pedersen, and J. A. Jensen. Clinical use and evaluation of coded excitation in B-mode images. In *Proc. IEEE Ultrason. Symp.*, volume 2, pages 1689–1693, 2000.

[7] T. X. Misaridis and J. A. Jensen. An effective coded excitation scheme based on a predistorted FM signal and an optimized digital filter. In *Proc. IEEE Ultrason. Symp.*, volume 2, pages 1589–1593, 1999.

[8] J. A. Jensen. Field: A program for simulating ultrasound systems. *Med. Biol. Eng. Comp.*, 10th Nordic-Baltic Conference on Biomedical Imaging, Vol. 4, Supplement 1, Part 1:351–353, 1996b.

Bibliography

[1] A. W. Rihaczek. *Principles of high-resolution radar*. McGraw-Hill, New York, 1969.

[2] J. A. Jensen. *Estimation of Blood Velocities Using Ultrasound: A Signal Processing Approach*. Cambridge University Press, New York, 1996.

[3] J. A. Jensen, O. Holm, L. J. Jensen, H. Bendsen, H. M. Pedersen, K. Salomonsen, J. Hansen, and S. Nikolov. Experimental ultrasound system for real-time synthetic imaging. In *Proc. IEEE Ultrason. Symp.*, volume 2, pages 1595–1599, 1999.

[4] K. K. Shung, M. B. Smith, and B. Tsui. *Principles of Medical Imaging*. Academic Press, New York, 1992.

[5] Gordon S. Kino. *Acoustic Waves, Devices, Imaging, & Analog Signal Processing*. Prentice-Hall, Englewood Cliffs, N.J., 1987.

[6] L. E. Kinsler, A. R. Frey, A. B. Coppens, and J. V. Sanders. *Fundamentals of Acoustics*. John Wiley & Sons, New York, third edition, 1982.

[7] V. A. Shutilov. *Fundamental Physics of Ultrasound*. Gordon and Breach, 1988.

[8] Peter N. T. Wells. *Biomedical Ultrasonics*. Academic Press, London, 1977.

[9] B. D. Steinberg. *Principles of aperture and array system design*. John Wiley & Sons, New York, 1976.

[10] Y. Takeuchi. An investigation of a spread energy method for medical ultrasound systems - Part one: theory and investigation. *Ultrasonics*, pages 175–182, 1979.

[11] M. O'Donnell. Coded excitation system for improving the penetration of real-time phased-array imaging systems. *IEEE Trans. Ultrason., Ferroelec., Freq. Contr.*, 39:341–351, 1992.

[12] R. Y. Chiao and L. J. Thomas. Synthetic transmit aperture imaging using orthogonal Golay coded excitation. In *Proc. IEEE Ultrason. Symp.*, pages 1677–1680, 2000.

[13] J. Y. Chapelon. Pseudo-random correlation imaging and system characterization. In *Progress in Medical Imaging*, pages 227–246, New York: Springer-Verlag, 1988. V. L. Newhouse.

[14] R. Kazys, L. Svilainis, and L. Mazeika. Application of orthogonal ultrasonic signals and binaural processing for imaging of the environment. *Ultrasonics*, 38:171–175, 2000.

[15] B. Haider, P. A. Lewin, and K. E. Thomenius. Pulse elongation and deconvolution filtering for medical ultrasonic imaging. *IEEE Trans. Ultrason., Ferroelec., Freq. Contr.*, 45:98–113, 1998.

[16] J. Shen and E. S. Ebbini. A new coded-excitation Ultrasound imaging system - Part I: Basic principles. *IEEE Trans. Ultrason., Ferroelec., Freq. Contr.*, 43:131–140, 1996.

[17] P. Li, E. Ebbini, and M. O'Donnell. A new filter design technique for coded excitation systems. *IEEE Trans. Ultrason., Ferroelec., Freq. Contr.*, 39:693–699, 1992.

[18] M. Pollakowski and H. Ermert. Chirp signal matching and signal power optimization in pulse-echo mode ultrasonic nondestructive testing. *IEEE Trans. Ultrason., Ferroelec., Freq. Contr.*, 41(5):655–659, 1994.

[19] N. A. H. K. Rao. Investigation of a pulse compression technique for medical ultrasound: a simulation study. *Med. Biol. Eng. Comp.*, 32(2):181–188, 1994.

[20] T. X. Misaridis and J. A. Jensen. An effective coded excitation scheme based on a predistorted FM signal and an optimized digital filter. In *Proc. IEEE Ultrason. Symp.*, volume 2, pages 1589–1593, 1999.

[21] C. Passmann and H. Ermert. A 100-MHz ultrasound imaging system for dermatologic and opthalmologic diagnostics. *IEEE Trans. Ultrason., Ferroelec., Freq. Contr.*, 43:545–552, 1996.

[22] A. V. Oppenheim and R. W. Schafer. *Discrete-Time Signal Processing*. Prentice-Hall, Englewood Cliffs, N.J., 1989.

[23] J. S. Bendat and A. G. Piersol. *Random Data. Analysis and Measurement Procedures*. John Wiley & Sons, New York, 2nd edition, 1986.

[24] H. L. van Trees. *Detection, estimation and modulation theory, part 1*. Wiley, New York, 1968.

[25] C. E. Cook and M. Bernfeld. *Radar signals*. Artech House, Inc., 1993.

[26] D. R. Wehner. *High-resolution radar*. Artech House, Boston, second edition, 1995.

[27] J. A. Jensen, D. Gandhi, and W. D. O'Brien. Ultrasound fields in an attenuating medium. In *Proc. IEEE Ultrason. Symp.*, pages 943–946, 1993.

[28] K. V. Gurumurthy and R. M. Arthur. A dispersive model for the propagation of ultrasound in soft tissue. *Ultrason. Imaging*, 4:355–377, 1982.

[29] A. W. Rihaczek. Radar waveform selection - a simplified approach. *IEEE Trans. Aerosp., Electron. Systems*, AES-7(6):1078–1086, 1971.

[30] J. C. Guey and M. R. Bell. Diversity waveform sets for delay-Doppler imaging. *IEEE Trans. Inform. Theroy*, 44(4):1504–1522, 1998.

[31] J. W. Adams. A new optimal window. *IEEE Trans. Sig. Pro.*, 39:1753–1769, 1991.

[32] M. R. Bell. Information theory and radar waveform design. *IEEE Trans. Inform. Theroy*, 39(5):1578–1597, 1993.

[33] J. D. Wolf, G. M. Lee, and C. E. Suyo. Radar waveform synthesis by mean-square optimization techniques. *IEEE Trans. Aerosp., Electron. Systems*, AES-5(4):611–619, 1969.

[34] R. L. Mitchell and A. W. Rihaczek. Matched-filter responses of the linear FM waveform. *IEEE Trans. Aerosp. Elect. Syst.*, AES-4(3):417–432, 1968.

[35] C. J. Oliver. *Synthetic aperture radar*. Artech House, Inc., 1998.

[36] M. I. Skolnik. *Radar Handbook*. McGraw-Hill, New York, second edition, 1990.

[37] F. J. Harris. On the use of windows for harmonic analysis with the discrete Fourier transform. *Proc. IEEE*, 66:51–83, 1978.

[38] M. Kowatsch and H. R. Stocker. Effect of Fresnel ripples on sidelobe suppression in low time-bandwidth product linear FM pulse compression. *IEE Proc.*, 129:41–44, 1982.

[39] A. W. Rihaczek and S. J. Hershkowitz. *Theory and practice of radar target identification*. Artech House, Boston, 2000.

[40] C. E. Cook and J. Paolillo. A pulse compression predistortion function for efficient sidelobe reduction in a high-power radar. *Proc. IEEE*, 52:377–389, 1964.

[41] D. V. Sarwate and M. B. Pursley. Crosscorrelation properties of pseudorandom and related sequences. *Proc. IEEE*, 68(5):593–619, 1980.

[42] D. Z. Dokovic. Equivalence classes and representatives of Golay sequences. *Discrete Math.*, 189:79–93, 1998.

[43] S. Z. Budisin. New complementary pairs of sequences. *Electr. Lett.*, 26(13):881–883, 1990.

[44] C. C. Tseng and C. I. Liu. Complementary sets of sequences. *IEEE Trans. Inform. Theroy*, IT-18(5):644–651, 1972.

[45] R. Sivaswamy. Multiphase complementary codes. *IEEE Trans. Inform. Theroy*, IT-24(5):546–552, 1978.

[46] R. Sivaswamy. Self-clutter cancellation and ambiguity properties of subcomplementary sequences. *IEEE Trans. Aerosp. Elect. Syst.*, AES-18(2):163–180, 1982.

[47] R. L. Frank. Polyphase codes with good nonperiodic correlation properties. *IEEE Trans. Inform. Theroy*, IT-9:43–45, 1963.

[48] B. L. Lewis and F. F. Kretschmer. A new class of polyphase pulse compression codes and techniques. *IEEE Trans. Aerosp. Elect. Syst.*, AES-17(3):364–372, 1981.

[49] B. M. Popovic. Generalized chirp-like polyphase sequences with optimum correlation properties. *IEEE Trans. Inform. Theroy*, 38(4):1406–1409, 1992.

[50] B. M. Popovic. Efficient matched filter for the generalized chirp-like polyphase sequences. *IEEE Trans. Aerosp. Elect. Syst.*, 30(3):769–777, 1994.

[51] K. R. Griep, J. A. Ritcey, and J. J. Burlingame. Poly-phase codes and optimal filters for multiple user ranging. *IEEE Trans. Aerosp. Elect. Syst.*, 31(2):752–767, 1995.

[52] J. M. Baden and M. N. Cohen. Optimal peak sidelobe filters for biphase pulse compression. In *IEEE Int. Radar Conf.*, pages 249–252, 1990.

[53] AIUM. Acoustic output measurements and labeling standard for diagnostic ultrasound equipment. Technical report, American Institute of Ultrasound in Medicine, Bethesda, Maryland, 1992.

[54] FDA. 510(k) guide for measuring and reporting acoustic output of diagnostic medical devices. Technical report, Center for Devices and Radiological Health, FDA, 1985.

[55] AIUM. Safety considerations for diagnostic ultrasound. Technical report, American Institute of Ultrasound in Medicine, Bethesda, Maryland, 1988.

[56] J. A. Jensen. Field: A program for simulating ultrasound systems. *Med. Biol. Eng. Comp.*, 10th Nordic-Baltic Conference on Biomedical Imaging, Vol. 4, Supplement 1, Part 1:351–353, 1996b.

[57] T. Misaridis and J. A. Jensen. User's guide for the AFG library for the LeCroy 9112 arbitrary function generator: MATLAB interface, version 1.1. Technical report, Center For Fast Ultrasound Imaging, Technical University of Denmark, 1999.

[58] C. E. Cook. Linear FM signal formats for beacon and communication systems. *IEEE Trans. Aerosp. Elect. Syst.*, AES-10(4):471–478, 1974.

[59] G. Chandran and J. S. Jaffe. Signal set design with constrained amplitude spectrum and specified time-bandwidth product. *IEEE Trans. Commun.*, 44(6):725–732, 1996.

[60] M. P. Lotter and L. P. Linde. A comparison of three families of spreading sequences for CDMA applications. *COMSIG-94*, pages 68–75, 1994.

[61] G. R. Lockwood, J. R. Talman, and S. S. Brunke. Real-time 3-D ultrasound imaging using sparse synthetic aperture beamforming. *IEEE Trans. Ultrason., Ferroelec., Freq. Contr.*, 45:980–988, 1998.

[62] C. R. Cooley and B. S. Robinson. Synthetic aperture imaging using partial datasets. In *Proc. IEEE Ultrason. Symp.*, pages 1539–1542, 1994.

[63] M. Karaman, P. C. Li, and M. O'Donnel. Synthetic aperture imaging for small scale systems. *IEEE Trans. Ultrason., Ferroelec., Freq. Contr.*, 42:429–442, 1995.

[64] R. Y. Chiao, L. J. Thomas, and S. D. Silverstein. Sparse array imaging with spatially-encoded transmits. In *Proc. IEEE Ultrason. Symp.*, pages 1679–1682, 1997.

[65] R. Y. Chiao and L. J. Thomas. Method and apparatus for ultrasonic synthetic transmit aperture imaging using orthogonal complementary codes. *US Patent*, 6,048,315, 2000.

[66] T. X. Misaridis, M. H. Pedersen, and J. A. Jensen. Clinical use and evaluation of coded excitation in B-mode images. In *Proc. IEEE Ultrason. Symp.*, volume 2, pages 1689–1693, 2000.

[67] T. Misaridis and J. A. Jensen. Space-time encoding for high frame rate ultrasound imaging. To be published. *Ultrasonics*, 2002.

[68] J. P. Costas. A study of a class of detection waveforms having nearly ideal range-Doppler ambiguity properties. *Proc. IEEE*, 72(8):996–1009, 1984.

[69] A. Freedman and N. Levanon. Staggered Costas signals. *IEEE Trans. Aerosp. Elect. Syst.*, 22(6):695–701, 1986.

[70] S. W. Golomb and H. Taylor. Constructions and properties of Costas arrays. *Proc. IEEE*, 72(9):1143–1163, 1984.

[71] J. P. Donohoe and F. M. Ingels. The ambiguity properties of FSK/PSK signals. *IEEE Int. Radar Conf.*, pages 268–273, 1990.

[72] J. S. Jaffe, D. Glassbrenner, and P. M. Cassereau. Incoherent coding techniques and performance characterization for multibeam sonar systems. In *IEEE Int. Conf. Acous., Speech, Sig. Pro.*, volume 5, pages 2709–2712, 1988.

[73] J. K. Hsiao and F. F. Kretschmer Jr. Design of a staggered p.r.f. moving target indication filter. *Radio, Electron. Engineer*, 43(11):689–693, 1973.

[74] H. Steyskal, R. A. Shore, and R. L. Haupt. Methods for null control and their effects on the radiation pattern. *IEEE Trans. Antenna Propag.*, AP-34(3):404–409, 1986.

[75] S. I. Nikolov, K. Gammelmark, and J. A. Jensen. Recursive ultrasound imaging. In *Proc. IEEE Ultrason. Symp.*, volume 2, pages 1621–1625, 1999.

Index